IBM Elastic Storage System 6.1.0.1

Quick Deployment Guide

SC28-3173-02

Note

Before using this information and the product it supports, read the information in "Notices" on page 77.

This edition applies to version 6 release 1 modification 0 of the following product and to all subsequent releases and modifications until otherwise indicated in new editions:

- IBM Spectrum® Scale Data Management Edition for IBM® ESS (product number 5765-DME)
- IBM Spectrum Scale Data Access Edition for IBM ESS (product number 5765-DAE)

IBM welcomes your comments; see the topic "How to submit your comments" on page xi. When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Figures

Tables

About this information

Who should read this information

This information is intended for administrators of IBM Elastic Storage® System (ESS) that includes IBM Spectrum Scale RAID.

IBM Elastic Storage System information units

IBM Elastic Storage System (ESS) 5000 documentation consists of the following information units.

Information unit	Type of information	Intended users
Hardware Guide	This unit provides ESS 5000 information including system overview, installing, and troubleshooting.	System administrators and IBM support team
Quick Deployment Guide	This unit provides ESS information including the software stack, deploying, upgrading, and best practices.	System administrators, analysts, installers, planners, and programmers of IBM Spectrum Scale clusters who are very experienced with the operating systems on which each IBM Spectrum Scale cluster is based
Model 092 storage enclosures	This unit provides information including initial hardware installation and setup, and removal and installation of field-replaceable units (FRUs), customer-replaceable units (CRUs) for ESS 5000 Expansion – Model 092, 5147-092.	System administrators and IBM support team
Model 106 storage enclosures	This unit provides information including hardware installation and maintenance for ESS 5000 Expansion – Model 106.	System administrators and IBM support team
Problem Determination Guide	This unit provides ESS 5000 information including setting up call home, replacing servers, issues, maintenance procedures, and troubleshooting.	System administrators and IBM support team
Command Reference	This unit provides information about ESS commands and scripts.	System administrators and IBM support team
IBM Spectrum Scale RAID: Administration	This unit provides IBM Spectrum Scale RAID information including administering, monitoring, commands, and scripts.	• System administrators of IBM Spectrum Scale systems • Application programmers who are experienced with IBM Spectrum Scale systems and familiar with the terminology and concepts in the XDSM standard

Related information

Related information

For information about:

- IBM Spectrum Scale, see:

  http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html
- mmvdisk command, see mmvdisk documentation.
- Mellanox OFED (MLNX_OFED v4.9-2.2.4.0) Release Notes, go to https://docs.mellanox.com/display/MLNXOFEDv492240/Release%20Notes
- DCS3700 storage enclosure, see:
  - *System Storage® DCS3700 Quick Start Guide*, GA32-0960-04:

    https://www-01.ibm.com/support/docview.wss?uid=ssg1S7005178
  - *IBM System Storage DCS3700 Storage Subsystem and DCS3700 Storage Subsystem with Performance Module Controllers: Installation, User's, and Maintenance Guide*, GA32-0959-07:

    http://www.ibm.com/support/docview.wss?uid=ssg1S7004920
- For information about the IBM Power Systems EXP24S I/O Drawer (FC 5887), see IBM Knowledge Center :

  http://www.ibm.com/support/knowledgecenter/8247-22L/p8ham/p8ham_5887_kickoff.htm
- IBM Spectrum Scale call home, see Understanding call home.
- Installing IBM Spectrum Scale and CES protocols with the installation toolkit, see Installing IBM Spectrum Scale on Linux® nodes with the installation toolkit.
- Detailed information about the IBM Spectrum Scale installation toolkit, see Using the installation toolkit to perform installation tasks: Explanations and examples.
- CES HDFS, see Adding CES HDFS nodes into the centralized file system.
- Installation toolkit ESS support, see ESS awareness with the installation toolkit.
- IBM POWER8® servers, see IBM Knowledge Center:

  http://www.ibm.com/support/knowledgecenter/POWER8/p8hdx/POWER8welcome.htm
- Extreme Cluster/Cloud Administration Toolkit (xCAT), go to the xCAT website :

  http://xcat.org/
  - xCAT 2.16.1 Release Notes®

For the latest support information about IBM Spectrum Scale RAID, see the IBM Spectrum Scale RAID FAQ in IBM Knowledge Center:

http://www.ibm.com/support/knowledgecenter/SSYSP8/gnrfaq.html

Conventions used in this information

Table 1 on page xi describes the typographic conventions used in this information. UNIX file name conventions are used throughout this information.

Table 1. Conventions

Convention	Usage
bold	Bold words or characters represent system elements that you must use literally, such as commands, flags, values, and selected menu options. Depending on the context, **bold** typeface sometimes represents path names, directories, or file names.
bold underlined	bold underlined keywords are defaults. These take effect if you do not specify a different keyword.
constant width	Examples and information that the system displays appear in constant-width typeface. Depending on the context, constant-width typeface sometimes represents path names, directories, or file names.
italic	*Italic* words or characters represent variable values that you must supply. *Italics* are also used for information unit titles, for the first use of a glossary term, and for general emphasis in text.
<*key*>	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, <Enter> refers to the key on your terminal or workstation that is labeled with the word *Enter*.
****	In command examples, a backslash indicates that the command or coding example continues on the next line. For example: `mkcondition -r IBM.FileSystem -e "PercentTotUsed > 90" \` `-E "PercentTotUsed < 85" -m p "FileSystem space used"`
{*item*}	Braces enclose a list from which you must choose an item in format and syntax descriptions.
[*item*]	Brackets enclose optional items in format and syntax descriptions.
<Ctrl-*x*>	The notation <Ctrl-*x*> indicates a control character sequence. For example, <Ctrl-c> means that you hold down the control key while pressing <c>.
***item*...**	Ellipses indicate that you can repeat the preceding item one or more times.
|	In *synopsis* statements, vertical lines separate a list of choices. In other words, a vertical line means *Or*. In the left margin of the document, vertical lines indicate technical changes to the information.

How to submit your comments

To contact the IBM Spectrum Scale development organization, send your comments to the following email address:

scale@us.ibm.com

Chapter 1. ESS Software deployment preparation

Install the ESS software package and deploy the storage servers by using the following information. The goal is to create a cluster allowing client or protocol nodes to access the file systems.

	ESS 3000	**ESS 5000**	**ESS Legacy**
Runs on	POWER8 or POWER9™ EMS	POWER9 EMS	POWER8 or POWER9 EMS
I/O node OS	Red Hat® Enterprise Linux 8.2 x86_64	Red Hat Enterprise Linux 8.2 PPC64LE	Red Hat Enterprise Linux 7.9 PPC64LE
Architecture	x86_64	PPC64LE	PPC64LE
IBM Spectrum Scale	5.1.0.3 efix8	5.1.0.3	5.1.0.3
OFED	MLNX_OFED_LINUX-4.9-2.2.4.0-rhel8.2-x86_64.iso Separate binary for firmware	MLNX_OFED_LINUX-4.9-2.2.5.1-rhel8.2-ppc64le.iso Firmware binary included	MLNX_OFED_LINUX-4.9-2.2.5.1 Firmware binary included
Firmware RPM	6009	6009	6009
SAS Adapter Firmware	N/A	16.00.11.00	16.00.11.00
Mpt3sas	N/A	34.00.00.00 (not in box)	34.00.00.00 (not in box)
Platform RPM	gpfs.ess.platform.ess3k-6.1.0-1.x86_64	N/A	N/A
Support RPM	gpfs.gnr.support-ess3000-1.0.0-2	gpfs.gnr.support-ess5000-1.0.0-1	gpfs.gnr.support-essbase-1.0.0-2.noarch.rpm
Podman	1.6.4 (1.4.4 RH7)	1.6.4	1.6.4 (1.4.4 RH7)
Container version	Red Hat Enterprise Linux 7.9	Red Hat Enterprise Linux 7.9	Red Hat Enterprise Linux 7.9
Ansible®	2.9.19	2.9.19	2.9.19
xCAT	2.16.1	2.16.1	2.16.1
PEMS	1111	N/A	N/A
Kernel	4.18.0-193.47.1.el8_2	4.18.0-193.47.1.el8_2	4.18.0-193.47.1.el8_2
POWER firmware	N/A	FW950.10(071_045) NVDIMM ver: Bundled BPM ver: Bundled	FW860.90 (SV860_226)
ndctl	N/A	ndctl-65-1.el8	N/A

	ESS 3000	ESS 5000	ESS Legacy
OPAL	N/A	opal-prd-ess.v4-1.el8.ppc64le.rpm	N/A
System firmware	Canister firmware FW 1111(2.02.000_0B0G_1.73_FB300052_0C32.official) Boot drive firmware • Smart -1236 • Micron -ML32	N/A	N/A
IPR	N/A	IPR 19512b00	IPR 19512b00
Enclosure firmware	N/A	5U92 - E558 4U106 - 5266	PPC64LE Slider 2U24 - 4230 5U84 - 4087 4U106 - 5266
NVMe firmware	SN1MSN1M	N/A	N/A
Network adapter	CX5CX5-VPI MT4120 = 16.28.2006 MT4121 = 16.28.2006 MT4122 = 16.28.2006 MT4123 = 20.28.2006 MT4125 = 22.28.2006	CX5HDR 100 MT4120 = 16.28.2006 MT4121 = 16.28.2006 MT4122 = 16.28.2006 MT4123 = 20.28.2006 MT4125 = 22.28.2006	MT4120 = 16.28.2006 MT4121 = 16.28.2006 MT4122 = 16.28.2006 MT4123 = 20.28.2006 MT4125 = 22.28.2006
ESA	ESA: esagent.pLinux-4.5.5-1	ESA: esagent.pLinux-4.5.5-1	ESA: esagent.pLinux-4.5.5-1
BIOS	52	N/A	N/A

Prerequisites

- This document (ESS Software Quick Deployment Guide)
- SSR completes physical hardware installation and code 20.
  - SSR uses Worldwide Customized Installation Instructions (WCII) for racking, cabling, and disk placement information.
  - SSR uses the respective ESS Hardware Guide (ESS 3000 or ESS 5000) for hardware checkout and setting IP addresses.
- Worksheet notes from the SSR
- Latest ESS tgz downloaded to the EMS node from Fix Central (If newer version is available).
  - Data Access Edition or Data Management Edition: Must match the order; if the edition does not match your order, open a ticket with the IBM Service.

- High-speed switch and cables have been run and configured.
- Low-speed host names are ready to be defined based on the IP addresses the SSR have configured.
- High-speed host names (suffix of low speed) and IP addresses are ready to be defined.
- Container host name and IP address are ready to be defined in the `/etc/hosts` file.
- Host and domain name (FQDN) are defined in the `/etc/hosts` file.
- **ESS Legacy 6.1.0.x Only:** You must convert to `mmvdisk` before deploying the ESS Legacy 6.1.0.x container. If you have not done so already, convert by `mmvdisk` by using the following steps:

  1. Check if there are any `mmvdisk` node classes.

     ```
     mmvdisk nodeclass list
     ```

     There should be one node class per ESS Legacy building-block. If the command output does not show `mmvdisk` for your ESS Legacy nodes, convert to `mmvdisk` before running the ESS Legacy 6.1.0.x container.
  2. Convert to `mmvdisk` by running the following command from one of the POWER8 IO nodes or from the POWER8 EMS node.

     ```
     gssgenclusterrgs -G gss_ppc64 --suffix=-hs --convert
     ```

     You can also use `-N` with a comma-separated list of nodes.

     Note: Wait for 5 minutes for daemons to recycle. The file system remains up.

What is in the /home/deploy directory on the EMS node

- ESS 5000 tgz used in manufacturing (may not be the latest)
- ESS 3000 tgz used in manufacturing (may not be the latest)
- ESS Legacy tgz used in manufacturing (may not be the latest)
- Red Hat Enterprise Linux 8.2 PPC64LE ISO (POWER9 EMS)
- Red Hat Enterprise Linux 7.9 PPC64LE ISO (POWER8 EMS)
  - This ISO is not needed for deployment but it is provided to restore the EMS node in case of a failure.

Support for signed RPMs

ESS or IBM Spectrum Scale RPMs are signed by IBM.

The PGP key is located in `/opt/ibm/ess/tools/conf`

```
-rw-r-xr-x 1 root root 907 Dec 1 07:45 SpectrumScale_public_key.pgp
```

You can check if an ESS or IBM Spectrum Scale RPM is signed by IBM as follows.

1. Import the PGP key.

   ```
   rpm --import  /opt/ibm/ess/tools/conf/SpectrumScale_public_key.pgp
   ```
2. Verify the RPM.

   ```
   rpm -K RPMFile
   ```

ESS 3000, ESS 5000, and ESS Legacy networking requirements

In any scenario you must have an EMS node and a management switch. The management switch must be split into 2 VLANS.

- Management VLAN
- Service/FSP VLAN

You also need a high-speed switch (IB or Ethernet) for cluster communication.

ESS 3000

POWER8 or POWER9 EMS

POWER9 EMS is preferred if it is a new ESS 3000 system without legacy (POWER8) building-blocks.

- If you are adding ESS 3000 to a POWER8 EMS:
  - An additional connection for the container to the management VLAN must be added. A C10-T2 cable must be run to this VLAN.
  - A public/campus connection is recommended in C10-T3.
  - A management connection must be run from C10-T1 (This should be already in place if adding to an existing POWER8 EMS with legacy nodes).
  - Port 1 on each ESS 3000 canister must be connected to the management VLAN.
- If you are using an ESS 3000 with a POWER9 EMS:
  - C11-T1 must be connected on the EMS to the management VLAN.
  - Port 1 on each ESS 3000 canister must be connected to the management VLAN.
  - C11-T2 must be connected on the EMS to the FSP VLAN.
  - HMC1 must be connected on the EMS to the FSP VLAN.

Note: It is highly recommended that you connect C11-T3 to a campus connection or run an additional management connection. If you do not do this step, you will lose the connection to the EMS node when the container starts.

ESS 5000

POWER9 EMS support only

EMS must have the following connections:

- C11-T1 to the management VLAN
- C11-T2 to the FSP VLAN
- HMC1 to the FSP VLAN

ESS 5000 nodes must have the following connections:

- C11-T1 to the management VLAN
- HMC1 to the FSP VLAN

Note: It is highly recommended that you connect C11-T3 to a campus connection or run an additional management connection. If you do not do this step, you will lose the connection to the EMS node when the container starts.

ESS Legacy

POWER8 or POWER9 EMS supported

POWER8 EMS must have the following connections:

- C10-T1 to the management VLAN
- C10-T4 to the FSP/Service VLAN
- C10-T2 to the management VLAN
- C10-T3 optional campus connection
- HMC1 to the FSP/Service VLAN

POWER9 EMS must have the following connections:

- C11-T1 to the management VLAN
- C11-T2 to the FSP VLAN

- HMC1 to the FSP VLAN
- C11-T3 to the campus or management network/VLAN

POWER8 nodes:

- C12-T1 to the management VLAN
- HMC1 to the FSP VLAN

Build naming conventions

There are three different releases in ESS 6.1.0.x, each with two editions: Data Management Edition and Data Access Edition. Example package names are as follows:

```
// Legacy
ess_legacy_6.1.0.1_0509-23_dme_ppc64le.tgz
ess_legacy_6.1.0.1_0509-23_dae_ppc64le.tgz

// ESS 5000
ess5000_6.1.0.1_0510-00_dme_ppc64le.tgz
ess5000_6.1.0.1_0510-00_dae_ppc64le.tgz

// ESS 3000
ess3000_6.1.0.1_0509-21_dme_ppc64le.tgz
ess3000_6.1.0.1_0509-21_dae_ppc64le.tgz
```

Note: The version shown here might not be the GA version available on IBM FixCentral. It is recommended to go to IBM FixCentral and download the latest code.

POWER8 considerations

If you are moving from an xCAT-based release (5.3.x) to a container based releases (6.1.x.x), the following considerations apply:

- You must add an additional management network connection to C10-T2.
- A public or additional management connection is recommended in C10-T3.
- You must stop and uninstall xCAT before installing the container.

POWER8 + POWER9 considerations

- If both POWER8 and POWER9 EMS nodes are in an environment, it is recommended that you use only the POWER9 EMS for management functions (containers, GUI, ESA, collector).
- Only a single instance of all management services is recommended and solely on the POWER9 EMS.
- POWER8 only needs to exist as a management node if you are mixing a non-container-based release (5.3.x) with a container-based release (6.x.x.x).
- It is recommended that all nodes in the storage cluster contain the same ESS release and IBM Spectrum Scale version.
- It is recommended that you upgrade to the latest level before adding a building block.

Note: If you are mixing ESS Legacy 5.3.x and ESS 3000 on a POWER8 EMS, the following considerations apply:

- You cannot upgrade the EMS node from the ESS 3000 container.
- ESS 3000 detects if xCAT is installed on the host EMS node. If xCAT is installed, it stops the upgrade.
- You must upgrade the EMS node by using the legacy deployment procedure outlined in *ESS 5.3.x Quick Deployment Guide*.

Migrating from an ESS Legacy environment (xCAT-based 5.3.x) to an ESS Legacy container-based environment (6.1.x.x)

The following guidance is for customers migrating from an xCAT-based release to a container-based release for POWER8 offerings.

POWER9 EMS

You cannot run both POWER8 and POWER9 EMS nodes in the same environment for ESS Legacy. If you are moving a POWER9 EMS, migrate all services from the POWER8 EMS and uninstall xCAT. You can then re-use the POWER8 EMS for other purposes such as quorum node, client node, or spare EMS. The preference is to always use a POWER9 EMS if possible and you must not run multiple instances of GUI, performance monitoring collectors, etc. in the same cluster. For this requirement, there are exceptions for certain stretch cluster environments and if you are mixing ESS Legacy and container-based deployments such as ESS 5.3.7 on POWER8 and ESS 6.0.2.x on POWER9.

POWER8 EMS

If you are migrating from ESS 5.3.x to ES 6.1.0.x on a POWER8 EMS, do the following steps.

1. Stop and uninstall x Cat by doing the following steps on a POWER8 EMS, outside of the container.
   a. Stop xCAT.

   ```
   systemctl stop xcatd
   ```

   b. Uninstall xCAT.

   ```
   yum remove xCAT*
   ```

   c. Remove dependencies.

   ```
   yum remove dbus-devel dhcp bind java-1.8.0-openjdk
   ```

2. Add container connection to C10-T2.
3. Update `/etc/hosts` with the desired container host name and IP address.

Other notes

- The following tasks must be complete before starting a new installation (tasks done by manufacturing and the SSR):
  - SSR has ensured all hardware is clean, and IP addresses are set and pinging over the proper networks (through the code 20 operation).
  - `/etc/hosts` is blank
  - The ESS tgz file (for the correct edition) is in the `/home/deploy` directory. If upgrade is needed, download from Fix Central and replace.
  - Network bridges are cleared.
  - Images and containers are removed.
  - SSH keys are cleaned up and regenerated.
  - All code levels are at the latest at time of manufacturing ship.
- Customer must make sure the high-speed connections are cabled and the switch is ready before starting.
- All node names and IP addresses in this document are examples.
- Changed root password should be same on each node, if possible. The default password is `ibmesscluster`. It is recommended to change the password after deployment is completed.
- Each server's IPMI and ASMI passwords (POWER® nodes only) are set to the server serial number. Consider changing these passwords when the deployment is complete.

ESS best practices

- ESS 6.x.x.x uses a new embedded license. It is important to know that installation of any Red Hat packages outside of the deployment upgrade flow is not supported. The container image provides everything required for a successful ESS deployment. If additional packages are needed, contact IBM for possible inclusion in future versions.
- For ESS 3000, consider enabling TRIM support. This is outlined in detail in *IBM Spectrum Scale RAID Administration*. By default, ESS 3000 only allocates 80% of available space. Consult with IBM development, if going beyond 80% makes sense for your environment, that is if you are not concerned about the performance implications due to this change.
- You must setup a campus or additional management connection before deploying the container.
- If running with a POWER8 and a POWER9 EMS in the same environment, it is best to move all containers to the POWER9 EMS. If there is a legacy PPC64LE system in the environment, it is best to migrate all nodes to ESS 6.1.x.x and decommission the POWER8 EMS altogether. This way you do not need to run multiple ESS GUI instances.
- If you have a POWER8 EMS, you must upgrade the EMS by using the legacy flow if there are xCAT based PPC64LE nodes in the environment (including protocol nodes). If there are just an ESS 3000 system and a POWER8 EMS, you can upgrade the EMS from the ESS 3000 container.
- If you are migrating the legacy nodes to ESS 6.1.x.x on the POWER8 EMS, you must first uninstall xCAT and all dependencies. It is best to migrate over to the POWER9 EMS if applicable.
- You must be at ESS 5.3.7 (Red Hat Enterprise Linux 7.7 / Python3) or later to run the ESS 3000 container on the POWER8 EMS.
- You must run the **`config load`** command against all the storage nodes (including EMS) in the cluster before enabling admin mode central or deploying the protocol nodes by using the installation toolkit.

Chapter 2. What's new and support matrix

The major changes and support matrix of this release are as follows.

Major changes from earlier releases

Release	Major changes
ESS 6.1.0.x • ESS Legacy (POWER8 PPC64LE) • ESS 3000 • ESS 5000	• Red Hat Enterprise Linux 8.2 • Red Hat Enterprise Linux 7.9 (ESS Legacy or POWER8 EMS) • IBM Spectrum Scale 5.1.0.3 • UBI 8 • CX6 VPI (ESS 5000) • Firmware 950 SP1 (ESS 5000)

Support matrix

Release	OS	Runs on	Can upgrade or deploy
ESS 3000 (6.1.0.x)	• Red Hat Enterprise Linux 7.9 (PPC64LE) • Red Hat Enterprise Linux 8.2 (x86_64)	• POWER8 EMS • POWER9 EMS	• ESS 3000 nodes • POWER8 EMS • POWER9 EMS • POWER8 protocol nodes • POWER9 protocol nodes
ESS 5000 (6.1.0.x)	• Red Hat Enterprise Linux 7.9 (PPC64LE)	• POWER9 EMS	• ESS 5000 nodes • POWER9 EMS • POWER9 protocol nodes
ESS Legacy (6.1.0.x)	• Red Hat Enterprise Linux 8.2 (PPC64LE) • Red Hat Enterprise Linux 7.9 (PPC64LE)	• POWER8 EMS • POWER9 EMS	• ESS POWER8 I/O nodes (PPC64LE) • ESS POWER8 protocol nodes (PPC64LE) • ESS POWER9 protocol nodes (PPC64LE)* * From POWER9 EMS only • POWER8 EMS • POWER9 EMS

Support notes, rules, and best practices

- Multiple EMS nodes are not supported in the same cluster. If you are adding a POWER9 EMS to an existing cluster run by a POWER8 EMS, the POWER9 EMS must be the only one used for management functions such as GUI, performance monitoring collector, etc.
- Multiple GUI instances are not supported in the same cluster.
- One collector node must be run at a time in the cluster. This must be on the same node as the GUI.
- You cannot mix major IBM Spectrum Scale versions in the storage cluster. All nodes must be updated to the latest level.
- ESA/Call home must be running on the EMS.
- If possible, run the client nodes in a separate cluster than the storage.
- If you are running a stretch cluster, you must ensure that each node has a unique `hostid`. The `hostid` might be non-unique if the same IP addresses and host names are being used on both sides of the stretch cluster. Run **gnrhealthcheck** before creating recovery groups when adding nodes in a stretch cluster environment. You can manually check the `hostid` on all nodes as follows:

  ```
  mmdsh -N { NodeClass | CommaSeparatedListofNodes } hostid
  ```

  If `hostid` on any node is not unique, you must fix by running **genhostid**. These steps must be done when creating a recovery group in a stretch cluster.

Chapter 3. ESS Common installation instructions

Note: You must convert to `mmvdisk` before using ESS Legacy 6.1.0.x.

The following common instructions need to be run for a new installation or an upgrade of an ESS system.

These instructions are based on steps required for a POWER9 EMS. Important POWER8 notes are outlined where needed. The following build is used for example purposes.

```
ess5000_6.1.0.1_0510-00_dme_ppc64le.tgz
```

Note: If you have protocol nodes, add them to the commands provided in these instructions. The default `/etc/hosts` file has host names `prt1` and `prt2` for protocol nodes. You might have more than two protocol nodes.

1. Log in to the EMS node by using the management IP (set up by SSR by using the provided worksheet). The default password is `ibmesscluster`.
2. **Set up a campus or a public connection (interface enP1p8s0f2).** Connect an Ethernet cable to C11-T3 on the EMS node to your lab network. This connection serves as a way to access the GUI or the ESA agent (call home) from outside of the management network. The container creates a bridge to the management network, thus having a campus connection is highly advised.

   Note: It is recommended but not mandatory to set up a campus or public connection. If you do not set up a campus or a public connection, you will temporarily lose your connection when the container bridge is created in a later step.

   This method is for configuring the campus network, not any other network in the EMS node. Do not modify T1, T2, or T4 connections in the system after they are set by SSR, and use the SSR method only to configure T1 and T2 (if changing is mandatory after SSR is finished). That includes renaming the interface, setting IP, or any other interaction with those interfaces.

   You can use the **`nmtui`** command to set the IP address of the campus interface. For more information, see Configuring IP networking with nmtui.
3. Complete the `/etc/hosts` file on the EMS node. This file must contain the low-speed (management) and high-speed (cluster) IP addresses, FQDNs, and short names. The high-speed names must contain a suffix to the low-speed names (For example, essio1-hs (high-speed name) to essio1 (low-speed name)). This file must also contain the container host name and the IP address.

   ```
   127.0.0.1 localhost localhost.localdomain.local localhost4 localhost4.localdomain4

   ## Management IPs 192.168.45.0/24
   192.168.45.20 ems1.localdomain.local ems1
   192.168.45.21 essio1.localdomain.local essio1
   192.168.45.22 essio2.localdomain.local essio2
   192.168.45.23 prt1.localdomain.local prt1
   192.168.45.24 prt2.localdomain.local prt2

   ## High-speed IPs 10.0.11.0/24
   10.0.11.1 ems1-hs.localdomain.local ems1-hs
   10.0.11.2 essio1-hs.localdomain.local essio1-hs
   10.0.11.3 essio2-hs.localdomain.local essio2-hs
   10.0.11.4 pr1-hs.localdomain.local prt1-hs
   10.0.11.5 pr2-hs.localdomain.local prt2-hs

   ## Container info 192.168.45.0/24
   192.168.45.80 cems0.localdomain.local cems0

   ## Protocol CES IPs
   10.0.11.100 prt_ces1.localdomain.local prt_ces1
   10.0.11.101 prt_ces1.localdomain.local prt_ces1
   10.0.11.102 prt_ces2.localdomain.local prt_ces2
   10.0.11.103 prt_ces2.localdomain.local prt_ces2
   ```

   Note:

- `localdomain.local` is just an example and cannot be used for deployment. You must change it to a valid fully qualified domain name (FQDN) during the `/etc/hosts` setup. The domain must be the same for each network subnet that is defined. Also, ensure that you set the domain on the EMS node (**hostnamectl set-hostname** *NAME*).

  NAME must be the FQDN of the management interface (T1) of the EMS node. If you need to set other names for campus, or other interfaces, those names must be the alias but not the main host name as returned by the **hostnamectl** command.

  You can set up the EMS FQDN manually or wait until prompted when the ESS deployment binary is started. At that time, the scripts confirms the FQDN and provides the user a chance to make changes.
- If you are planning to set up an ESS 3000 or ESS Legacy PPC64LE system with the ESS 5000 EMS node, add the ESS 3000 host names to `/etc/hosts` by using the same structure (low-speed (management) and high-speed (cluster) IP addresses, FQDNs, and short names).
- Do not use any special characters, underscores, or dashes in the host names other than the high speed suffix (example: -hs). Doing this might cause issues with the deployment procedure.

4. Clean up the old containers and images.

   Note: Typically, this is applicable only for upgrades.

   a. List the containers.

   ```
   podman ps -a
   ```

   b. Stop and remove the containers.

   ```
   podman stop ContainerName
   podman rm ContainerName -f
   ```

   c. List the images.

   ```
   podman images
   ```

   d. Remove the images.

   ```
   podman image rm ImageID -f
   ```

   e. [Recommended] Remove container bridges as follows.

      i) List the currently configured bridges.

      ```
      nmcli c
      ```

      ii) Clean up any existing bridges before the new container is set up. The bridge names must be `mgmt_bridge` and `fsp_bridge`.

      ```
      nmcli c del BridgeName
      ```

5. Do additional clean up.

   - If you are using a POWER8 EMS and converting from the xCAT-based deployment to container, you must first stop and uninstall xCAT as follows.

     ```
     systemctl stop xcatd
     yum -y remove xCAT
     ```

   - Make sure that the DHCP server is not longer running.

     ```
     ps -ef | grep -i dhcp
     ```

   - Stop the GUI temporarily until upgrade or conversion from xCAT deployment to container is complete.

```
systemctl stop gpfsgui
```

6. Extract the installation package.

   Note: Ensure that you check the version that is installed from manufacturing (SSR worksheet). If there is a newer version available on Fix Central, replace the existing image in `/home/deploy` with the new image and then remove the old tgz file before doing this step.

```
cd /home/deploy

tar zxvf ess5000_6.1.0.1_0510-00_dme_ppc64le.tgz

ess5000_6.1.0.1_0510-00_dme_ppc64le.sh
ess5000_6.1.0.1_0510-00_dme_ppc64le.sh.sha256
```

7. Accept the license and install the accepted image.

```
./ess5000_6.1.0.1_0510-00_dme_ppc64le.sh --start-container
```

   During this step, you are first prompted to accept the license agreement. Press 1 to accept. You are then prompted to input answers to 3 questions before the installation starts (2 questions for ESS 3000).

   - Confirm or set EMS FQDN.
   - Provide the container short name.
   - Provide a free IP address on the FSP subnet for the container FSP connection. (Not applicable to ESS 3000)

   Example of contents of the extracted installation package:

```
ess5000_6.1.0.1_0510-00_dme_ppc64le.dir/
ess5000_6.1.0.1_0510-00_dme.dir/ess5000_6.1.0.1_0510-00_dme_ppc64le.tar
ess5000_6.1.0.1_0510-00_dme.dir/ess5000_6.1.0.1_0510-00_dme_ppc64le_binaries.iso
ess5000_6.1.0.1_0510-00_dme.dir/rhel-8.2-server-ppc64le.iso
ess5000_6.1.0.1_0510-00_dme.dir/podman_rh7.tgz
ess5000_6.1.0.1_0510-00_dme.dir/podman_rh8.tgz
ess5000_6.1.0.1_0510-00_dme.dir/Release_note.ess5000_6.1.0.1_0510-00_dme_ppc64le.txt
ess5000_6.1.0.1_0510-00_dme.dir/python3-site-packages_rh7.tgz
ess5000_6.1.0.1_0510-00_dme.dir/python3_rh7.tgz
ess5000_6.1.0.1_0510-00_dme.dir/Release_note.ess5000_6.1.0.1_0510-00_dme_ppc64le.txt
ess5000_6.1.0.1_0510-00_dme.dir/essmkyml
ess5000_6.1.0.1_0510-00_dme.dir/essmgr
ess5000_6.1.0.1_0510-00_dme.dir/essmgr_p8.yml
ess5000_6.1.0.1_0510-00_dme.dir/essmgr_p9.yml
ess5000_6.1.0.1_0510-00_dme.dir/data/
ess5000_6.1.0.1_0510-00_dme.dir/classes/
ess5000_6.1.0.1_0510-00_dme.dir/logs/
ess5000_6.1.0.1_0510-00_dme.dir/essmgr.yml
```

   In this step, you are prompted to provide these inputs:

   - Container name (must be in `/etc/hosts` or be resolvable by using DNS)
   - Container FSP IP address (must be on the same network block that is set on C11-T2)
   - Confirmation of the EMS FQDN (must match what is set for the management IP in `/etc/hosts`). If this value needs to be changed or set, **`essmkyml`** helps with that task. **`essmkyml`** is located in the extracted directory (example: `/home/deploy/ess5000_6.0.2.1_0508-00_dme.dir/`)
   - EMS host name must be on the management network (also called xCAT). Other networks can be aliases (A) or canonical names (CNAME) on DNS or on the `/etc/hosts` file.

```
Is the current EMS FQDN c145f05zems06.gpfs.net correct (y/n):
```

   - Remember not to add the DNS domain `localdomain` to the input:

```
Please type the desired and resolvable short hostname [ess5k-cems0]: cems0
```

- Remember that the IP address must belong to the 10.0.0.x/24 network block (It is assumed that the recommended FSP network was used):

```
Please type the FSP IP of the container [10.0.0.5]: 10.0.0.80
```

Note: The values in parentheses ([ ]) are just examples or the last entered values.

If all of the checks pass, the `essmgr.yml` file is written and you can proceed to bridge creation, if applicable, and running the container.

Note: If you are deploying ESS 3000, you are not prompted to answer the FSP IP question.

At this point, if all checks are successful, the image is loaded and container is started. Example:

```
ESS 5000 CONTAINER root@cems0:/ #
```

8. Run the config load command. The **`config load`** command determines the node information based on VPD and also exchange the SSH keys.

```
essrun -N essio1,essio2,ems1 config load -p ibmesscluster
```

Note:

- Always include the EMS in this command along with all nodes of the same type in the building-blocks.
- Use the low-speed management host names. Specify the root password with `-p`.
- The password (-p) is the root password of the node. By default, it is `ibmesscluster`. Consider changing the root password after deployment is complete.
- To verify the groups that are created during the **`config load`** step, run **`lsdef -t group`**.

After this command is run, you can use `-G` for future **`essrun`** steps (For example, `-G ess_ppc64le`). There are different node group names for ESS 3000 and ESS Legacy. Use **`lsdef -t group`** to view the other options if applicable.

Chapter 4. ESS New deployment instructions

Before starting with these steps, you must complete the steps in Chapter 3, “ESS Common installation instructions,” on page 11.

The following steps are covered in this topic:

- Upgrading the EMS and IO nodes, if required.
- Creating network bonds.
- Creating the cluster.
- Adding the EMS node to the cluster.
- Creating the file system.
- Configuring performance monitoring and starting the GUI.
- Setting up call home.
- Setting up time server.
- Final health checks.

Note: You can update by using the management node names (management) or after the **`config load`** is run, you can update by using a group of nodes. The groups are as follows:

- PPC64LE - ESS 5000 and ESS Legacy: `ess_ppc64le`
- x86_64 - ESS 3000: `ess_x86_64`

When the group is referenced in these instructions, `ess_ppc64le` is used as an example. If you are in an ESS 3000 environment, use `ess_x86_64`.

For the EMS node, you can use the group `ems`.

At this point, the user has already determined if an upgrade is required. If the version initially found in `/home/deploy` on the EMS node is earlier than the latest available on IBM Fix Central, the latest version should be already downloaded and deployed according to Chapter 3, “ESS Common installation instructions,” on page 11.

1. If an upgrade is required, upgrade the EMS node.

```
essrun -N ems1 update --offline

Please enter 'accept' indicating that you want to update the following list of nodes: ems1
>>> accept
```

Note:

- If the kernel is changed, you are prompted to leave the container, reboot the EMS node, restart the container, and run this command again.

  For example:

```
essrun -N ems1 --offline
Exit
systemctl reboot
```

  Navigate back to ESS 6.0.2.x extracted directory and run the following commands:

```
./essmgr -r
essrun -N ems1 --offline
```

- You cannot upgrade a POWER8 EMS currently running ESS Legacy code (5.3.x with xCAT control) from an ESS 3000 container. If xCAT is installed on the host, you must first uninstall it and cleanup any dependencies before attempting an EMS upgrade from the container. Do not remove xCAT if legacy deployment is not needed, typically only if you are moving to ESS Legacy 6.1.0.x container. If

you are still using an ESS Legacy deployment (5.3.x), update the EMS by using the upgrade instructions outlined in *ESS 5.3.x Quick Deployment Guide*.

2. If required, update the IO nodes.

```
essrun -G ess_ppc64le update --offline
```

3. Create network bonds.

```
essrun -G ess_ppc64le network --suffix=-hs
essrun -N ems1 network --suffix=-hs
```

4. Run the network test.

   This test uses **nsdperf** to determine if the newly created network bonds are healthy.

   SSH from the container to an I/O node or the EMS node.

```
ssh essio1
ESSENV=TEST essnettest -N essio1,essio2 --suffix=-hs
```

   This command performs the test with an optional RDMA test afterward if there is Infiniband. Ensure that there are no errors in the output indicating dropped packets have exceeded thresholds. When completed, type `exit` to return back to the container.

5. Create the cluster.

```
essrun -G ess_ppc64le cluster --suffix=-hs
```

6. Add the EMS node to the cluster.

```
essrun -N essio1 cluster --add-ems ems1 --suffix=-hs
```

7. Create the file system.

```
essrun -G ess_ppc64le filesystem --suffix=-hs
```

   Note:

   - By default, this command attempts to use all the available space. If you need to create multiple file systems or a CES shared root file system for protocol nodes, consider using less space. For example:

```
essrun -G ess_ppc64le filesystem --suffix=-hs --size 80%
```

   - This step creates combined metadata + data vdisk sets by using a default RAID code and block size. You can use additional flags to customize or use the **mmvdisk** command directly for advanced configurations.
   - If you are updating ESS 3000, the default set-size is 80% and it must not be increased. For additional options, see *essrun command*. The default block size for PPC64LE is 16M whereas for ESS 3000 it is 4M.
   - If you are deploying protocol nodes, make sure that you leave space for CES shared root file system. Adjust the set-size slightly lower when you are creating this required file system for protocol nodes.

Final setup instructions

1. From the EMS node (outside of the container), configure and start the performance monitoring collector.

```
mmperfmon config generate --collectors ems1-hs
```

2. From the EMS node (outside of the container), configure and start the performance monitoring sensors.

```
mmchnode --perfmon -N ems1-hs,essio1-hs,essio2-hs
```

3. Capacity and fileset quota monitoring is not enabled in the GUI by default. You must correctly update the values and restrict collection to the EMS node only.

   a. To modify the GPFS Disk Capacity collection interval, run the following command.

   ```
   mmperfmon config update GPFSDiskCap.restrict=EMSNodeName
            GPFSDiskCap.period=PeriodInSeconds
   ```

   The recommended period is 86400 so that the collection is done once per day.

   b. To restrict GPFS Fileset Quota to run on the management server node only, run the following command.

   ```
   mmperfmon config update GPFSFilesetQuota.period=600 GPFSFilesetQuota.restrict=EMSNodeName
   ```

   Here the *EMSNodeName* must be the name shown in the **mmlscluster** output.

   Note: To enable quota, the filesystem quota checking must be enabled. Refer **mmchfs -Q** and **mmcheckquota** commands in *IBM Spectrum Scale: Command and Programming Reference*.

4. Verify that the values are set correctly in the performance monitoring configuration by running the **mmperfmon config show** command on the EMS node. Ensure that `GPFSDiskCap.period` is properly set, and `GPFSFilesetQuota` and `GPFSDiskCap` are both restricted to the EMS only.

   Note: If you are moving from manual configuration to auto configuration then all sensors are set to default. Make the necessary changes using the **mmperfmon** command to customize your environment accordingly. For information on how to configure various sensors using **mmperfmon**, see Manually installing IBM Spectrum Scale GUI.

5. Start the performance collector on the EMS node.

   ```
   systemctl start pmcollector
   ```

6. Start the GUI.

   ```
   systemctl start gpfsgui
   ```

   a. Create the GUI admin user.

   ```
   /usr/lpp/mmfs/gui/cli/mkuser UserName -g SecurityAdmin
   ```

   b. In a web browser, enter the public or campus IP address with `https` and walk through the System Setup wizard instructions.

7. Log in to each node and run the following command.

   ```
   essinstallcheck -N localhost
   ```

   Doing this step verifies that all software and cluster versions are up-to-date.

8. From the EMS node, outside of the container, run the following final health check commands to verify your system health.

   ```
   gnrhealthcheck
   mmhealth node show -a
   ```

9. Set the time zone and set up Chrony.

   Before getting started, ensure that Chrony and time zone are set correctly on the EMS and I/O nodes. Refer to Appendix D, “How to set up chronyd (time server),” on page 49 to perform these tasks before proceeding.

10. Set up call home. For more information, see Appendix B, “Configuring call home in ESS 5000, ESS 3000, and ESS Legacy,” on page 29.

    The supported call home configurations are:

    - Software call home

- Node call home (including for protocol nodes)
- Drive call home

11. Refer to Appendix G, “Client node tuning recommendations,” on page 57.

Chapter 5. ESS Upgrade instructions

Warning: You must have a clean and healthy system before starting any ESS upgrade (online or offline). At least, the following commands must run free of errors when run on any node outside of container:

```
gnrhealthcheck
mmhealth node show -a
```

You can also run the **essrun healthcheck** command instead, from inside the container.

```
essrun -G NodeGroup healthcheck
```

Upgrade can be done by using the following methods

- Offline upgrade: This method requires a given node or nodes to have GPFS shut down before beginning. This method is faster than online update, in which nodes are upgraded in parallel including firmware, but the system is typically taken down for a period of time.
- Online upgrade: This method allows the cluster to stay fully available and the code is typically updated one node per building-block in parallel.

Note:

- The EMS node and protocol node upgrades are available only in the offline mode.
- You must update a given node type from its associated container. For example, update ESS 5000 nodes from the ESS 5000 container. Refer to the matrix for protocol node and EMS support.

Online upgrade assumptions (IO nodes only):

- The cluster is created with EMS, one or more ESS nodes, and optionally one or more ESS building blocks or protocol nodes.
- The file system is built and recovery groups are active and healthy.
- GPFS is active on all ESS nodes and quorum is achieved.
- New container is installed that will update the code on the EMS and I/O nodes.
- GUI and collector services are stopped on the EMS before starting the upgrade.

Before starting the online upgrade, make sure that all ESS nodes are active by running the following command from one of the cluster nodes:

```
mmgetstate -N NodeClass
```

Where *NodeClass* is your ESS 3000, ESS 5000, or ESS Legacy node class. For more information, see *mmlsnodeclass command*.

Offline upgrade assumptions (EMS or protocol nodes only):

- You assume the risks of potential quorum loss.
- The GPFS GUI and collector must be down.

Note: Before upgrading the protocol nodes, consult the IBM Spectrum Scale toolkit documentation. You might need to shut down services on a given protocol node before the upgrade can start.

1. Complete the steps in Chapter 3, “ESS Common installation instructions,” on page 11. Make sure that you add the protocol nodes to the configuration load if you are planning to upgrade protocol nodes.
2. Update the EMS node first.

```
essrun -N ems1 update --offline
```

If kernel version changed during the update, you are prompted to exit the container, reboot, rerun the container, and rerun the update command.

```
Seems that kernel has changed. This will require a reboot
Please exit container and reboot ems1
Restart container (./essmgr -r) once ems1 is back and run update again.
```

After the reboot and restarting the container, run the EMS node update again.

```
essrun -N ems1 update --offline
```

Note: You cannot upgrade a POWER8 EMS currently running ESS Legacy code (5.3.x with xCAT control) from an ESS 3000 container. If xCAT is installed on the host, you must first uninstall it and cleanup any dependencies before attempting an EMS upgrade from the container. Do not remove xCAT if legacy deployment is not needed, typically only if you are moving to ESS Legacy 6.1.0.x container. If you are still using an ESS Legacy deployment (5.3.x), update the EMS by using the upgrade instructions outlined in *ESS 5.3.x Quick Deployment Guide*.

3. Update the protocol nodes.

```
essrun -N prt01,prt02 update --offline
```

4. Run installation check on each node type by logging in to EMS node and protocol nodes.

```
essinstallcheck
```

5. Do ESS IO nodes offline update as follows.

Important: For doing an offline update, GPFS must be down in the ESS cluster. The GPFS status is checked. If it is up on a given node, you are asked if it is OK to shut it down.

If you want to do an online update of IO nodes, refer to "Update ESS IO nodes online" on page 21.

- Update by using the group of all configured ESS nodes.

```
essrun -G ess5k_ppc64le update --offline
```

- Update by using the individual nodes.

```
essrun -N essio1,essio2 update --offline
```

- Update one node at a time.

```
essrun -N essio1 update --offline
```

These command examples show ESS 5000 node and node classes, but you can use these commands with ESS 3000 and ESS Legacy nodes and node classes as well.

After offline update is done, proceed to starting GPFS on the nodes.

6. Run installation check on each node from outside the container.

```
essinstallcheck
```

7. Start GPFS on all nodes.

```
mmstartup -N NodeList | NodeGroup
```

Wait for a few minutes and then check the state.

```
mmgetstate -N NodeList | NodeGroup
mmgetstate -s
```

Note: If any protocol nodes are updated, ensure that you restart CES services on those nodes.

Update ESS IO nodes online

1. Update the IO nodes online by using one of the following commands.

   Important: For doing an online upgrade, recovery groups must be correctly created in both I/O nodes from the ESS cluster. Quorum is checked early in the process. If no quorum is achieved, the upgrade stops.

   - Update by using the group of all configured ESS IO nodes.

     ```
     essrun -G ess_ppc64le update
     ```

   - Update by using the individual nodes.

     ```
     essrun -N essio1,essio2 update
     ```

2. Run installation check on each updated node.

   ```
   essinstallcheck
   ```

3. Change the **autoload** parameter to enable GPFS to automatically start on all nodes.

   ```
   mmchconfig autoload=yes
   ```

Final steps for online and offline upgrade

Do the following final steps after the online or the offline update is complete.

1. Start the performance monitoring collector on the EMS node.

   ```
   systemctl start pmcollector
   ```

2. Start the performance monitoring sensors on each node.

   ```
   mmdsh -N NodeList | NodeGroup "systemctl restart pmsensors"
   ```

3. Start the GUI on the EMS node.

   ```
   systemctl start gpfsgui
   ```

4. Run manual health checks on each node.

   ```
   gnrhealthcheck
   mmhealth node show -a
   ```

5. Start the ESA GUI (call home), if applicable.

Appendix A. ESS known issues

Known issues in ESS

The following table describes the known issues in IBM Elastic Storage System (ESS) and how to resolve these issues.

Issue	Resolution or action	Product
JAVA_HOME might be pointing to the wrong version which might cause ESA startup to fail: In the following example, note how Java™ is pointing to the wrong location. This causes the ESA startup to fail: `# ls -alt` `total 20` `drwxr-xr-x.  2 root root 4096 Nov 22 15:02 .` `lrwxrwxrwx  1 root root 62 Nov 22 15:02` **`java -> /usr/lib/jvm/ java-11-openjdk-11.0.ea.28-7.el7.ppc64le/bin/java`** `lrwxrwxrwx  1 root root 70 Nov 22 15:02 java.1.gz -> /usr/share/man/ man1/java-java-11-openjdk-11.0.ea.28-7.el7.ppc64le.1.gz` `lrwxrwxrwx  1 root root 61 Nov 22 15:02 jjs -> /usr/lib/jvm /java-11-openjdk-11.0.ea.28-7.el7.ppc64le/bin/jjs`	To fix the problem, remove the current `java` symbolic link, update the `java` pointer, and retry the ESA activation. 1. Remove the current `java` symbolic link. `# cd /etc/alternatives/` `# rm java` `rm: remove symbolic link ‘java’? y` 2. Update the `java` pointer. `# ln -s /usr/lpp/mmfs/java java` `# ls -alt | grep -i java` `lrwxrwxrwx  1 root root  18 Nov 22 16:03 java -> /usr/lpp/mmfs/java` `cd /opt/ibm/` `# ln -s /etc/alternatives/java java-ppc64le-80` `# ls -alt` `total 0` `drwxr-xr-x.  5 root  root  62 Nov 22 16:04 .` `lrwxrwxrwx  1 root  root  22 Nov 22 16:04 java-ppc64le-80 -> /etc/alternatives/java` `dr-xr-x---  12 root  root  151 Nov 22 15:48 esa` `drwxr-xr-x. 10 root  root  119 Nov  7 16:09 ..` `drwx------  8 scalemgmt scalemgmt 121 Nov  7 16:00 wlp` `drwxr-xr-x.  7 root  root  68 Nov  7 14:36 gss` `# vi /opt/ibm/esa/runtime/conf/javaHome.sh` `# cat /opt/ibm/esa/runtime/conf/javaHome.sh` `JAVA_HOME=/opt/ibm/java-ppc64le-80/jre` 3. Retry the ESA activation. `# /opt/ibm/esa/bin/activator -C -p 5024 -w -Y`	ESS 3000
The hardware CPU validation GPFS callback is only active for one node in the cluster. This callback prevents GPFS from starting if a CPU socket is missing.	No action is required.	ESS 3000
During rolling upgrade, **`mmhealth`** might show the error `local_exported_ fs_unavail` even though the file system is still mounted.	During a rolling upgrade (Updating of one ESS I/O node at a time but maintaining quorum), **`mmhealth`** might display an error indicating that the local exported file system is unavailable. This message is erroneous.	• ESS 3000 • ESS 5000

Issue	Resolution or action	Product
	`` Component Status Status Change Reasons --- -------------- GPFS HEALTHY 6 min. ago - NETWORK HEALTHY 20 min. ago - FILESYSTEM DEGRADED 18 min. ago local_exported_fs_unavail(gpfs1) DISK HEALTHY 6 min. ago - NATIVE_RAID HEALTHY 6 min. ago - PERFMON HEALTHY 19 min. ago - THRESHOLD HEALTHY 20 min. ago - ``` The workaround is to restart **mmsysmon** on each node called out by **mmhealth**.	
During upgrade, if the container had an unintended loss of connection with the target canister(s), there might be a timeout of up to 2 hours in the Ansible update task.	Wait for the timeout and retry the **essrun** update task.	ESS 3000
During storage MES upgrade, you are required to update the drive firmware to complete the task. Some of the drives might not update on the first pass of running the command.	Re-run the **mmchfirmware –type drive** command which should resolve the issue and update the remaining drives.	ESS 3000
When running **essrun** commands, you might see messages such as these: ``` Thursday 16 April 2020 20:52:44 +0000 (0:00:00.572) 0:13:19.792 ******** Thursday 16 April 2020 20:52:45 +0000 (0:00:00.575) 0:13:20.367 ******** Thursday 16 April 2020 20:52:46 +0000 (0:00:00.577) 0:13:20.944 ******** Thursday 16 April 2020 20:52:46 +0000 (0:00:00.576) 0:13:21.521 ******** Thursday 16 April 2020 20:52:47 +0000 (0:00:00.570) 0:13:22.091 ******** Thursday 16 April 2020 20:52:47 +0000 (0:00:00.571) 0:13:22.663 ******** ```	This is a restriction in the Ansible timestamp module. It shows timestamps even for the “skipped” tasks. If you want to remove timestamps from the output, change the `ansible.cfg` file inside the container as follows: 1. vim `/etc/ansible/ansible.cfg` 2. Remove `,profile_tasks` on line 7. 3. Save and quit: `esc + :wq`	• ESS 3000 • ESS 5000
When running the **essrun config load** command, you might see a failure such as this: ``` stderr: |- rc=2 code=186 Failed to obtain the enclosure device ```	This failure means that the pems module is not running the canister. For fixing this, do the following: 1. Log in to the failed canister and run the following commands: ``` cd /install/ess/otherpkgs/rhels8/x86_64/gpfs yum reinstall gpfs.ess.platform.ess3k* ```	ESS 3000

Issue	Resolution or action	Product
`name with rc=2` `rc=2 code=669`	2. When the installation finishes, wait until the **lsmod | grep pems** command returns output similar to this: `pemsmod 188416 0` `scsi_transport_sas 45056 1 pemsmod` 3. Retry the **essrun config load** command from the container.	
Running **essrun -N node1,node2,… config load** command with high-speed names causes issues with the upgrade task using the `-G` flag.	The **essrun config load** command is an Ansible wrapper that attempts to discover the ESS 3000 canister node positions, place them into groups, and fix the SSH keys between the servers. This command must always be run using the low-speed or management names. You must not use the high-speed names with this command. For example: **essrun -N ess3k1a,ess3k1b config load** If you run this command using the high-speed or cluster names, this might result in issues when performing the update task. Example of what not to do: **essrun -N ess3k1a-hs,ess3k1b-hs config load** To confirm that the config run is set up correctly, use the **lsdef** command. This command returns only the low-speed or management names defined in `/etc/hosts`.	• ESS 3000 • ESS 5000
After reboot of an ESS 5000 node, systemd could be loaded incorrectly. Users might see the following error when trying to start GPFS: `Failed to activate service 'org.freedesktop.systemd1': timed out`	Power off the system and then power it on again. 1. Run the following command from the container: **rpower** *NodeName* **off** 2. Wait for at least 30 seconds and run the following command to verify that the system is off: **rpower** *NodeName* **status** 3. Restart the system with the following command. **rpower** *NodeName* **on**	ESS 5000
In ESS 5000 SLx series, after pulling a hard drive out for a long time wherein the drive has finished draining, when you re-insert the drive, the drive could not be recovered.	Run the following command from EMS or IO node to revive the drive: **mmvdisk pdisk change --rg** *RGName* **--pdisk** *PdiskName* **--revive** Where *RGName* is the recovery group that the drive belongs to and *PdiskName* is the drive's pdisk name.	ESS 5000
After the deployment is complete, if firmware on the enclosure, drive, or HBA adapter does not match the expected level, and if you run **essinstallcheck**, the following	The error about **mmvdisk** settings can be ignored. The resolution is to update the mismatched firmware levels on enclosure, adapter, or HBA adapters to the correct levels. You can run the **mmvdisk configuration check** command to confirm. List the **mmvdisk** node classes: **mmvdisk nc list**	• ESS 3000 • ESS 5000

Issue	Resolution or action	Product
mmvdisk settings related error message is displayed: `[ERROR] mmvdisk settings do NOT match best practices. Run mmvdisk server configure --verify --node-class ess5k_ppc64le_mmvdisk to debug.`	**Note: essinstallcheck** detects inconsistencies from **mmvdisk** best practices for all node classes in the cluster and stops immediately if an issue is found.	
When running **essinstallcheck** you might see an error message similar to: `System Firmware could not be obtained which will lead to a false-positive PASS message when the script completes.`	Rerun **essinstallcheck** which should properly query the firmware level.	ESS 5000
When running the **essrun - N** *Node* **healthcheck** command, the **essinstallcheck** script might fail due to incorrect error verification which might lead to an impression that there is a problem where there is none.	This health check command (**essrun - N** *Node* **healthcheck**) is removed from the ESS documentation and it is advised to use the manual commands to verify system health after deployment. Run the following commands for health check: • **gnrhealthcheck** • **mmhealth node show -a** • **essinstallcheck -N localhost**: This command needs to be run on each node.	• ESS 3000 • ESS 5000
During command-less disk replacement, there is a limit on how many disks can be replaced at one time.	For command-less disk replacement using commands, only replace up to 2 disks at a time. If command-less disk replacement is enabled, and more than 2 disks are replaceable, replace the 1st 2 disks, and then use the commands to replace the 3rd and subsequent disks.	• ESS 3000 • ESS 5000
Issue reported with command-less disk replacement warning LEDs.	The replaceable disk will have the amber led on, but not blinking. Disk replacement should still succeed.	ESS 5000
After upgrading an ESS 3000 node to version 6.1.0.x, the `pmsensors` service needs to be manually started.	After the ESS 3000 upgrade is complete, the `pmsensors` service does not automatically start. You must manually start the service for performance monitoring to be restored. On each ESS 3000 canister, run the following command: `systemctl start pmsensors` For checking the status of the service, run the following command: `systemctl status --no-pager pmsensors`	ESS 3000
ESS commands such as **essstoragequickcheck**, **essinstallcheck** must be run using **-N localhost**. If using the hostname such as **-N ess3k1a**, an error occurs.	There is currently an issue with running the ESS deployment commands by using the hostname of a node. The workaround is to run checks locally on each node by using localhost. For example, instead of using **essstoragequickcheck -N ess3k1a**, use the following command: `essstoragequickcheck -N localhost`	• ESS 3000 • ESS 5000

Issue	Resolution or action	Product
Hyperthreading might be enabled on an ESS 3000 system due to an incorrect kernel grub flag being set.	Hyperthreading needs to be disabled on ESS 3000 systems. This is ensured in following ways: • Disabled in BIOS • Disabled using the tuned profile • Disabled using the grub command line When disabled with the grub command line, the issue occurs because the grub configuration had an incorrect flag set in earlier versions. To resolve this issue, do the following: 1. Edit the `/etc/grub2.cfg` file to change `nohup` with `nosmt`. Before change: `set default_kernelopts="root=UUID=9a4a93b8-2e6b-4ba6-bda4-a7f8c3cb908f ro nvme.sgl_threshold=0 sshd=1 pcie_ports=native` **`nohup`** `resume=UUID=c939121b-526a-4d44-8d33-693f2fb7f018 rd.md.uuid=f6dbf6f2:8ac82ed6:875ca663:0094ac11 rd.md.uuid=06c2d5b0:c6603a1e:5df4b4d3:98fd5adc rhgb quiet crashkernel=4096M"` After change: `set default_kernelopts="root=UUID=9a4a93b8-2e6b-4ba6-bda4-a7f8c3cb908f ro nvme.sgl_threshold=0 sshd=1 pcie_ports=native` **`nosmt`** `resume=UUID=c939121b-526a-4d44-8d33-693f2fb7f018 rd.md.uuid=f6dbf6f2:8ac82ed6:875ca663:0094ac11 rd.md.uuid=06c2d5b0:c6603a1e:5df4b4d3:98fd5adc rhgb quiet crashkernel=4096M"` 2. Reboot the node for the changes to take effect.	ESS 3000
GUI Hardware page displays FN1 Enclosure state failed. However, unidentified reason of failed state.	There is currently no workaround for this issue.	ESS Legacy
Hardware details: All end points are not visible.	There is currently no workaround for this issue.	ESS Legacy
[ERROR] Network adapter MT4115 firmware: found 12.28.2006 expected 12.27.2008	Ignore this message. The correct version is 12.28.2006.	ESS Legacy
In an existing cluster with quorum nodes not exceeding 7 nodes, addition of any new nodes fails irrespective of the firmware level.	This is not considered a problem thus no workaround is needed.	• ESS 3000 • ESS 5000 • ESS Legacy
`essinstallcheck` might throw an error with garbage value: `Unsupported Adapter found HBA %25209405W%2D8i8e`	There is currently no workaround for this issue.	ESS Legacy

Issue	Resolution or action	Product
Running **essinstallcheck -N Ionode1 --phy-mapping** throws an exception.	There is currently no workaround for this issue.	ESS Legacy
When you run **mmlsenclosure all** on IO nodes, the following message gets displayed only for 2U24 enclosures, which are PPC64BE: `No enclosures were found`	There is currently no workaround for this issue.	ESS Legacy
When enabling security, the following error occurs: `ERROR! We were unable to read either as JSON nor YAML, these are the errors we got from each:` `JSON: Expecting value: line 1 column 1 (char 0)` `Syntax Error while loading YAML.` `could not find expected ':'` `The error appears to be in '/opt/ibm/ess/deploy/ansible/roles/security/tasks/securitydisable.yml': line 131, column 3,` `but may be elsewhere in the file depending on the exact syntax problem.`	To enable the security tool, change the line 130 in the `/opt/ibm/ess/deploy/ansible/roles/security/tasks/securitydisable.yml` file as follows: Replace `when not portmapperdown.stat.exists` with `when: portmapperdown.stat.exists != True` **Note:** There are two spaces before `when:` Code snippet before the change: `with_items:` `- systemctl start rpcbind > /dev/null 2>&1` `- systemctl start rpcbind.socket > /dev/null 2>&1` `when  not portmapperdown.stat.exists` Code snippet after the change: `with_items:` `- systemctl start rpcbind > /dev/null 2>&1` `- systemctl start rpcbind.socket > /dev/null 2>&1` `when: portmapperdown.stat.exists != True`	• ESS 3000 • ESS 5000 • ESS Legacy

Appendix B. Configuring call home in ESS 5000, ESS 3000, and ESS Legacy

In ESS 5000, ESS 3000, and ESS Legacy systems, ESS version 6.1.0.x can generate call home events when a drive in an attached enclosure needs to be replaced. ESS 5000, ESS 3000, and ESS Legacy can also generate call home events for other hardware-related events in the I/O server nodes, protocol nodes, and client nodes that need service.

ESS 5000 and ESS Legacy hardware events rely on POWER system OPAL logs and ESS 3000 hardware events rely on **mmcallhome** and **mmhealth** commands.

ESS version 6.1.x automatically opens an IBM Service Request with service data, such as the location and field replaceable unit (FRU) number to carry out the service task.

Disk call home for ESS 5000, ESS 3000, and ESS Legacy

The IBM Spectrum Scale RAID pdisk is an abstraction of a physical disk. A pdisk corresponds to exactly one physical disk, and belongs to exactly one de-clustered array within exactly one recovery group.

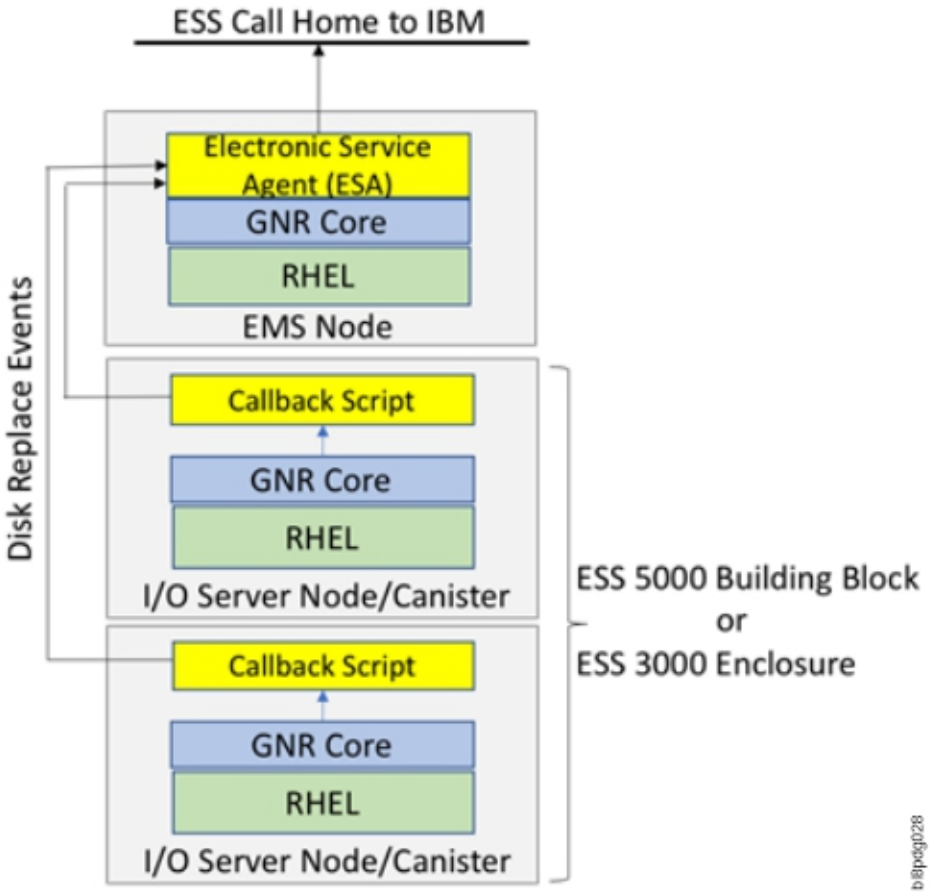


Figure 1. ESS Call Home Block Diagram

The attributes of a pdisk includes the following:

- The state of the pdisk
- The disk's unique worldwide name (WWN)
- The disk's field replaceable unit (FRU) code

- The disk's physical location code

When the pdisk state is ok, the pdisk is healthy and functioning normally. When the pdisk is in a diagnosing state, the IBM Spectrum Scale RAID disk hospital is performing a diagnosis task after an error has occurred.

The disk hospital is a key feature of the IBM Spectrum Scale RAID that asynchronously diagnoses errors and faults in the storage subsystem. When the pdisk is in a missing state, it indicates that the IBM Spectrum Scale RAID is unable to communicate with a disk. If a missing disk becomes reconnected and functions properly, its state changes back to ok. For a complete list of pdisk states and further information on pdisk configuration and administration, see IBM Spectrum Scale RAID Administration.

Any pdisk that is in the dead, missing, failing or slow state is known as a non-functioning pdisk.

When the disk hospital concludes that a disk is no longer operating effectively and the number of non-functioning pdisks reaches or exceeds the replacement threshold of their de-clustered array, the disk hospital adds the replace flag to the pdisk state. The replace flag indicates the physical disk corresponding to the pdisk that must be replaced as soon as possible. When the pdisk state becomes replace, the drive replacement callback script is run.

The callback script communicates with the ESA over a REST API. The ESA is installed in the ESS as part of the Management Server (EMS).

The EMS node initiates a call home task. The ESA is responsible for automatically opening a Service Request (PMR) with IBM support, and managing the end-to-end life cycle of the problem.

Installing the IBM Electronic Service Agent

IBM Electronic Service Agent (ESA) for PowerLinux version 4.5.5.1 or later can monitor the ESS systems. ESA is pre-installed on the EMS node when the EMS node is shipped.

The esagent rpm is also provided in the ESS 5000 or ESS 3000 or ESS Legacy binaries.iso file in the container package. The ISO is mounted when **essmgr** is run to start the container. When mounted, the rpm file can be found at the following location:

- ESS 5000: /install/ess/otherpkgs/rhels8/ppc64le/ess/
- ESS 3000: /install/ess/otherpkgs/rhels8/x86_64/ess/
- ESS Legacy: /install/ess/otherpkgs/rhels7/ppc64le/ess/

Issue the following command to verify that the rpm for the esagent is installed:

```
rpm -qa | grep esagent
```

This gives an output similar to the following:

```
esagent.pLinux-4.5.5-1.noarch.rpm
```

The RPM should be installed during manufacturing. In case it is not installed, issue the following command:

```
cd /install/ess/otherpkgs/rhels8/ppc64le/ess/
yum install esagent.pLinux-4.5.5-1.noarch.rpm
```

Login, activation, and configuration of ESA

After ESA is installed, it must be activated by using the **/opt/ibm/esa/bin/activator -C** command. Then, the ESA portal can be reached by going to the following link.

```
https://<EMS or ip>:5024/esa
```

For example:

```
https://192.168.45.20:5024/esa
```

ESA uses port 5024 by default. It can be changed by using the ESA CLI if needed. For more information on ESA, see IBM Electronic Service Agent. On the Welcome page, log in to the IBM Electronic Service Agent GUI. If an untrusted site certificate warning is received, accept the certificate or click **Yes** to proceed to the IBM Electronic Service Agent GUI. You can get the context sensitive help by selecting the **Help** option located in the upper right corner.

After you have logged in, go to the **Main Activate ESA**, to run the activation wizard. The activation wizard requires valid contact, location and connectivity information.

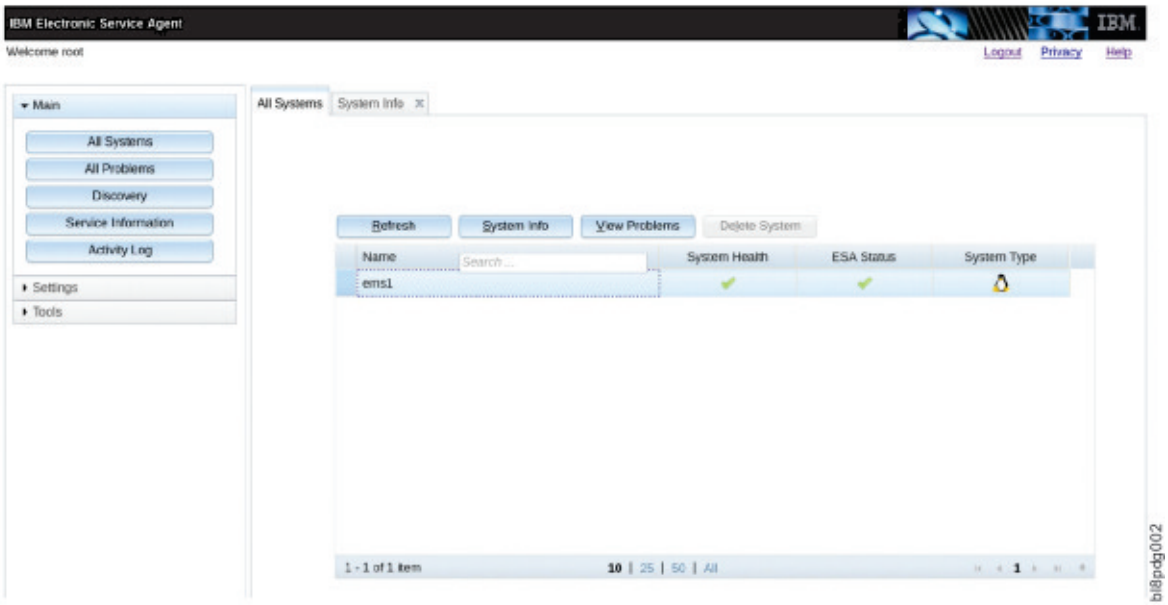


Figure 2. ESA portal after login

The `All Systems` menu option shows the node where ESA is installed. For example, ems1. The node where ESA is installed is shown as `PrimarySystem` in the **System Info**. The ESA status is shown as **Online** only on the `PrimarySystem` node in the **System Info** tab.

Note: ESA is not activated by default. In case it is not activated, you will get a message similar to the following message:

```
[root@ems1 tmp]# esscallhomeconf -E ems1 --show
E] IBM Electronic Service Agent (ESA) is not activated.
[I] Activated ESA using /opt/ibm/esa/bin/activator -C and provide customer detail via ESA GUI
later and retry.
[I] Alternatively use --esa-config switch and provide all customer details to do ESA activation
from here only.
[I] See --esa-config switch for further CLI activation of ESA
Exiting...
```

The **esscallhomeconf** has a new switch called **--esa-config**. Earlier users could activate ESA by using the **/opt/ibm/esa/bin/activator -C** command and defer the ESA configuration with the ESA GUI (earlier method of activating ESA before 6.x). However, with the introduction of **--esa-config** and its supporting switch, the activation of ESA and its configuration can also be done by using the **esscallhomeconf** command.

This switch can be used to activate ESA and configure ESA by using the CLI with the required customer information such as customer name, email ID, server location, etc. The user earlier provided this information when activating ESA by using the ESA GUI.

The usage information of **esscallhomeconf** is as follows.

```
usage: esscallhomeconf [-h] -E ESA-AGENT [--prefix PREFIX] [--suffix SUFFIX]
                       [--verbose] [--esa-hostname-fqdn ESA_HOSTNAME_FQDN]
                       [--stop-auto-event-report] [-N NODE-LIST] [--show]
                       [--register {node,all}] [--no-swcallhome] [--icn ICN]
                       [--serial SOLN-SERIAL] [--model SOLN-MODEL]
                       [--esa-config] [-m ESA_CONFIG_M]
```

```
                       [-u ESA_CONFIG_U] [-n ESA_CONFIG_N]
                       [-e ESA_CONFIG_E] [-t ESA_CONFIG_T]
                       [-f ESA_CONFIG_F] [-j ESA_CONFIG_J]
                       [-k ESA_CONFIG_K] [-g ESA_CONFIG_G]
                       [-a ESA_CONFIG_A] [-z ESA_CONFIG_Z]
                       [-y ESA_CONFIG_Y] [-r ESA_CONFIG_R]
                       [-b ESA_CONFIG_B] [-s ESA_CONFIG_S]
                       [-i ESA_CONFIG_I] [-p ESA_CONFIG_P] [-w] [-Y]
optional arguments:
  -h, --help            show this help message and exit
  -E ESA-AGENT          Provide nodename for esa agent node
  --prefix PREFIX       Provide hostname prefix. Use = between --prefix and
                        value if the value starts with -.
  --suffix SUFFIX       Provide hostname suffix. Use = between --suffix and
                        value if the value starts with -.
  --verbose             Provide verbose output
  --esa-hostname-fqdn ESA_HOSTNAME_FQDN
                        Fully qualified domain name of ESA server for
                        certificate validation.
  --stop-auto-event-report
                        Stop report of automatic event to ESA in case of any
                        hardware call home event reported to system.
  -N NODE-LIST          Provide a list of nodes to configure.
  --show                Show call home configuration details.
  --register {node,all}
                        Register endpoints(nodes, enclosure or all) with ESA.
  --no-swcallhome       Do not configure software callhome while configuring
                        hardware callhome
  --icn ICN             Provide IBM Customer Number for Software callhome.
  --serial SOLN-SERIAL  Provide ESS solution serial number.
  --model SOLN-MODEL    Provide ESS model. Applicable only for BE (ppc64)
                        models.
  --esa-config          Provide info for configuration of ESA via CLI.
  -m ESA_CONFIG_M       name of organization that owns or is responsible for
                        this system
  -u ESA_CONFIG_U       country or region where the system is located
  -n ESA_CONFIG_N       name of the primary person in your organization who
is responsible for this system
  -e ESA_CONFIG_E       email address for the primary contact person (e.g.
                        myuserid@mycompany.com)
  -t ESA_CONFIG_T       telephone number where the primary contact person can be reached
  -f ESA_CONFIG_F    secondary person in your organization who is
                     responsible for this system
  -j ESA_CONFIG_J    secondary person email address (e.g.
                     myuserid@mycompany.com)
  -k ESA_CONFIG_K    secondary person telephone number where the person can
                     be reached
  -g ESA_CONFIG_G    country or region of the contact person
  -a ESA_CONFIG_A    state or province where the system is located
  -z ESA_CONFIG_Z    postal code where the system is located
  -y ESA_CONFIG_Y    city where the system is located
  -r ESA_CONFIG_R    address where the system is located
  -b ESA_CONFIG_B    building where the system is located
  -s ESA_CONFIG_S    telephone number where the system is located
  -i ESA_CONFIG_I    IBM ID
  -p ESA_CONFIG_P    port number on which the subsystem listens for
                     incoming client requests. Default: 5024
  -w                 Add firewall rules to access ESA UI from remote
                     systems. Default: True
  -Y                 accept license agreement without displaying it.
                     Default: False
```

There are several switches which start with ESA_CONFIG that can be used with the **--esa-config** switch of the **esscallhomeconf** command to activate ESA by using the CLI instead of using the ESA GUI and activating it.

Entities or systems that can generate events are called endpoints. The EMS, I/O server nodes, and attached enclosures can be endpoints in ESS. Servers and enclosure endpoints can generate events. Server can generate hardware events which could be CPU, DIMM, OS Disk, etc. Typically, these events are also logged in the OPAL log. Enclosure generated call home is mostly during the disk replacement event.

In ESS, ESA is only installed on the EMS node, and it automatically discovers the EMS as `PrimarySystem`. The EMS node and I/O server nodes must be registered to ESA as endpoints.

The **esscallhomeconf** command is used to perform the registration task. The command also registers enclosures attached to the I/O servers by default.

Software call home can also be registered based on the customer information given while configuring the ESA agent. A software call home group `auto` is configured by default and the EMS node acts as the software call home server. The weekly and daily software call home data collection configuration is also activated by default. The software call home uses the ESA network connection settings to upload the data to IBM. The ESA agent network setup must be complete and working for the software call home to work.

Attention: You can configure software call home without running the **esscallhomeconf** command on the ESS system by using the **mmcallhome** command. However, it is recommended to not enable software call home with **mmcallhome** on any of the ESS systems including ESS 3000, ESS 5000, and ESS Legacy 5.x systems.

A sample output of the **esscallhomeconf** command is as follows.

```
# esscallhomeconf -E essems1 -N essems1,essio1,essio2 --esa-config
--register all --icn 123456789 -m IBMTEST -u INDIA -n UserName -e username@example.com
-t ContactNum -f UserName -j username@example.com -k ContactNum -g INDIA -a State
-z PostalCode -y Location -r Address -b Building -s ContactNum -i IBMTest -Y --crvpd --serial
212867A
--model 8247-21L

[I] ESA is activated but the configuration was not done.
[I] Activating ESA via CLI using information provided by --esa-config switch
[I] Successfully activated the ESA with customer detail...
2021-02-18T09:41:18.190176 Generating node list...
2021-02-18T09:41:35.966228 nodelist:  essems1   essio1  essio2
Existing vpd file found. --crvpd option is ignored.
End point essems1 registered successfully with systemid 1dab83cc3b9409d5bbf6e657c7e312c8
End point essio1 registered successfully with systemid f7e01a43e9a7464da6cfbe757ca9a669
End point essio2 registered successfully with systemid 1438fddb414738cf60dcade90570059
Skipping node gssems1 as it's not an IO node. Only IO nodes are attached to enclosures. Thus
only IO nodes are eligible to be registered their enclosures here.
End point enclosure G51704M registered successfully with systemid
e2b14722f6940b1c410c6ec4452ded9d
End point enclosure G517022 registered successfully with systemid
506794d52b9fd5f7c580ca8e48a051cc
ESA configuration for ESS Call home is complete.
Started configuring software callhome
Checking for ESA is activated or not before continuing.
Fetching customer detail from ESA.
Customer detail has been successfully fetched from ESA.
Setting software callhome customer detail.
Successfully set the customer detail for software callhome.
Enabled daily schedule for software callhome.
Enabled weekly schedule for software callhome.
Direct connection will be used for software calhome.
Successfully set the direct connection settings for software callhome.
Enabled software callhome capability.
Creating callhome automatic group
Created auto group for software call home and enabled it.
Software callhome configuration completed.
```

After running this single command, ESA is activated and configured, and the nodes are registered along with the enclosures. Software call home is also set up with the same command.

Configuring only hardware call home and skipping software call home configuration

The software call home feature collects files, logs, traces, and details of certain system health events from different nodes and services in an IBM Spectrum Scale cluster.

These details are shared with the IBM support center for monitoring and problem determination. For more information on call home, see Installing call home and Understanding call home.

You can configure hardware call home only by using the **esscallhomeconf** command. Use the `--no-swcallhome` option to set up just the call home hardware, and skip the software call home setup.

```
# /opt/ibm/ess/tools/bin/esscallhomeconf -N ems4,essio41,essio42 --suffix=-ce
-E ems4 --register all --no-swcallhome -m IBM -n UserName -e username@example.com
-t ContactNum -g "India" -s ContactNum -u "India" -r Address
-y Location -a State -z PostalCode -b Building
-f UserName -j username@example.com -k ContactNum -p 5024 -w -Y -i username@example.com --esa-
config
```

```
[I] ESA is activated but the configuration was not done.
[I] Activating ESA via CLI using information provided by --esa-config switch
[I] Successfully activated the ESA with customer detail...
2021-03-01T23:33:53.840111 Generating node list...
2021-03-01T23:34:02.716569 nodelist:  ems4      essio41 essio42
2021-03-01T23:34:02.716638 suffix used for endpoint hostname: -ce
End point ems4-ce registered successfully with systemid 6304ce01ebe6dfb956627e90ae2cb912
End point essio41-ce registered successfully with systemid a575bdce45efcfdd49aa0b9702b22ab9
End point essio42-ce registered successfully with systemid 5ad0ba8d31795a4fb5b327fd92ad860c
Skipping node ems4 as it's not an IO node. Only IO nodes are attached to enclosures. Thus only
IO nodes are eligible to be registered their enclosures here.
End point enclosure 78ZA006 registered successfully with systemid
32eb1da04b60c8dbc1aaaa9b0bd74976
ESA configuration for ESS Call home is complete.
Skipping software callhome configuration.
```

Attention: For ESS 3000 system, do not use the **--no-swcallhome** switch otherwise the ESS 3000 hardware call home will also not function.

Note: For ESS 5000 or ESS Legacy systems, you can skip the software call home functionality by using the **--no-swcallhome** switch if you do not want to use the software call home function.

Note: If any software only call home configuration was done earlier by using **mmcallhome** command or by using **esscallhomeconf** then subsequent execution of **esscallhomeconf** with **--no-swcallhome** does not clear the earlier configuration of software call home configuration. A software call home configuration that exists from before on an ESS cluster can be overwritten only if **esscallhomeconf** command is executed on the node without the **--no-swcallhome** switch. If you use **--no-swcallhome** with **esscallhomeconf** command and want to keep any older software call home configuration, it might cause issues or affect the ESS 3000 hardware call home functionality. Therefore, in any case it is advisable to clean up an earlier software call home configuration or overwrite the software call home configuration before configuring hardware and software call home function on ESS nodes by using the **esscallhomeconf** command.

Note: Re-running of the **esscallhomeconf** checks if ESA has been activated and its configuration was already done earlier or not. If yes then ESA configuration must be cleared before re-running **esscallhomeconf** by using **/opt/ibm/esa/bin/unconfig.sh**.

Attention: ESS Server Node Registration Entitlement in 6.1.x.x: With the introduction of the ESS 5000 and ESS 3000, the node registration entitlement of ESS node has changed. For earlier version of ESS, the ESS building blocks are registered as a solution MTMs. Usually the MTM of the EMS node is used as an entire building block Solution MTM. However, in case of ESS 5000 and ESS 3000, the system is registered based on its actual server serial number and MTM instead of solution MTM. Therefore, it is very important to check with IBM that the nodes are registered at the entitlement server before configuring hardware call home. ESS Legacy will continue to use the solution-based MTM registration and entitlement even in ESS 6.1.x.x.

ESS call home logs and location

The **esscallhomeconf** command logs the progress and error messages in the `/var/log/messages` file and the `/var/log/ess/` folder.

There is a **--verbose** option that provides more details of the progress and error messages. The following example displays the type of information sent to the `/var/log/messages` file on the EMS node by the **esscallhomeconf** command.

```
# grep essems4 /var/log/messages | grep esscallhomeconf
Feb 23 10:07:14 essems4 /esscallhomeconf: End point ems1-ib registered successfully with
systemid ed28297131f0d2b469edffc505a9708c
Feb 23 10:07:20 essems4 /esscallhomeconf: [I] End point essio11-ib registered successfully with
systemid db25a7e21ff4298243078806f964c495
Feb 23 10:07:20 essems4 /esscallhomeconf: [I] End point essio12-ib registered successfully with
systemid f9887f2bba6dee858146206dda96eb48
Feb 23 10:07:51 essems4 /esscallhomeconf: [I] End point proto11-ib registered successfully with
systemid 12ca060f30f9276cd52828fc117b0675
Feb 23 10:26:59 essems1 /esscallhomeconf: [I] End point enclosure EB15089 registered
successfully with systemid 72fadb281627047372f9ada47ed2fcb4
Feb 23 10:27:05 essems1 /esscallhomeconf: [I] End point enclosure EB15094 registered
successfully with systemid b266c524642846255f38a493e99bf10a
```

```
Feb 23 10:27:05 essems1 /esscallhomeconf: [I] End point enclosure EB15090 registered
successfully with systemid 8f9a45df3eb6137f6890ab18cf4c2957
End point enclosure EB15090
Feb 23 10:27:28 essems1 /esscallhomeconf: [I] ESA configuration for ESS Call home is complete.
Feb 23 10:28:04 essems1 /esscallhomeconf: [I] Software callhome configuration completed.
```

Attention: The **esscallhomeconf** command also configures the IBM Spectrum Scale call home setup. The IBM Spectrum Scale call home feature collects files, logs, traces, and details of certain system health events from the I/O and EMS nodes and services running on those nodes. These details are shared with the IBM support center for monitoring and problem determination. For more information on IBM Spectrum Scale call home, see IBM Spectrum Scale documentation in IBM Documentation.

Note: The ESS 3000 hardware call home is backed by software call home. In other words, software call home must be configured by using the **esscallhomeconf** command, without the **--no-swcallhome** switch in the ESS 3000 environment. Otherwise, the ESS 3000 hardware failure events are not reported to ESA and a PMR does not get opened.

The endpoints are visible in the ESA portal after registration, as shown in the following figure:

Figure 3. ESA portal after node registration

Name

Shows the name of the endpoints that are discovered or registered.

SystemHealth

Shows the health of the discovered endpoints. A green icon (√) indicates that the discovered system is working fine. The red (X) icon indicates that the discovered endpoint has some problem.

ESAStatus

Shows that the endpoint is reachable. It is updated whenever there is a communication between the ESA and the endpoint.

SystemType

Shows the type of system being used. Following are the various ESS device types that the ESA supports.

ESS Device type	Icon
ESS Application	
Disk	
Disk Enclosure	
Management Server	
Node	
Physical Server	
Virtual Server	
Other	

Figure 4. List of icons showing various ESS device types

Detailed information about the node can be obtained by selecting **System Information**. Here is an example of the system information:

System Information

Property	Value
Name	essio12.isst.gpfs.ibm.net
Machine Type	8247
Machine Model	22L
Serial Number	2145B3A
Manufacturer	IBM
Operating System	Linux
OS Type	Linux
OS Version	3.10.0-327.36.3.el7.ppc64
OS Additional Version	
IP Address	192.168.1.103 192.168.2.103
Firmware	
PM Enabled	No
ESA Status	Offline
System ID	898fb33e04f5ea12f2f5c7ec0f8516d4

Figure 5. System information details

When an endpoint is successfully registered, the ESA assigns a unique system identification (system id) to the endpoint. The system ID can be viewed using the **--show** option.
For example:

```
# esscallhomeconf -E ems4 --show
System id and system name from ESA agent
{
    "32eb1da04b60c8dbc1aaaa9b0bd74976": "78ZA006",
    "6304ce01ebe6dfb956627e90ae2cb912": "ems4-ce",
    "a575bdce45efcfdd49aa0b9702b22ab9": "essio41-ce",
```

```
    "5ad0ba8d31795a4fb5b327fd92ad860c": "essio42-ce"
}
```

When an event is generated by an endpoint, the node associated with the endpoint must provide the system id of the endpoint as part of the event. The ESA then assigns a unique event id for the event. The system id of the endpoints are stored in a file called `esaepinfo01.json` in the `/vpd` directory of the EMS and I/O servers that are registered. The following example displays a typical `esaepinfo01.json` file:

```
# cat /vpd/esaepinfo01.json
{
    "encl": {
        "78ZA006": "32eb1da04b60c8dbc1aaaa9b0bd74976"
    },
    "esaagent": "ems4",
    "node": {
        "ems4-ce": "6304ce01ebe6dfb956627e90ae2cb912",
        "essio41-ce": "a575bdce45efcfdd49aa0b9702b22ab9",
        "essio42-ce": "5ad0ba8d31795a4fb5b327fd92ad860c"
    }
}
```

The endpoints are visible in the ESA portal after registration. For more information, see IBM Spectrum Scale call home documentation.

Overview of a problem report

After ESA is activated, and the endpoints for the nodes and enclosures are registered, they can send an event request to ESA to initiate a call home.

For example, when `replace` is added to a pdisk state, indicating that the corresponding physical drive needs to be replaced, an event request is sent to ESA with the associated system id of the enclosure where the physical drive resides. After ESA receives the request it generates a call home event. Each server in the ESS is configured to enable callback for IBM Spectrum Scale RAID related events. These callbacks are configured during the cluster creation, and updated during the code upgrade. ESA can filter out duplicate events when event requests are generated from different nodes for the same physical drive. ESA returns an event identification value when the event is successfully processed. The ESA portal updates the status of the endpoints. The following figure shows the status of the enclosures when the enclosure contains one or more physical drives identified for replacement.

ESS 5000, ESS 3000 and ESS Legacy Disk enclosure failure call home event

- `pdReplacePdisk` - When a pDisk needs replacement inside an enclosure, ESA will get an event and a corresponding PMR will get opened. This event is not meant to report the server operating disk failure.

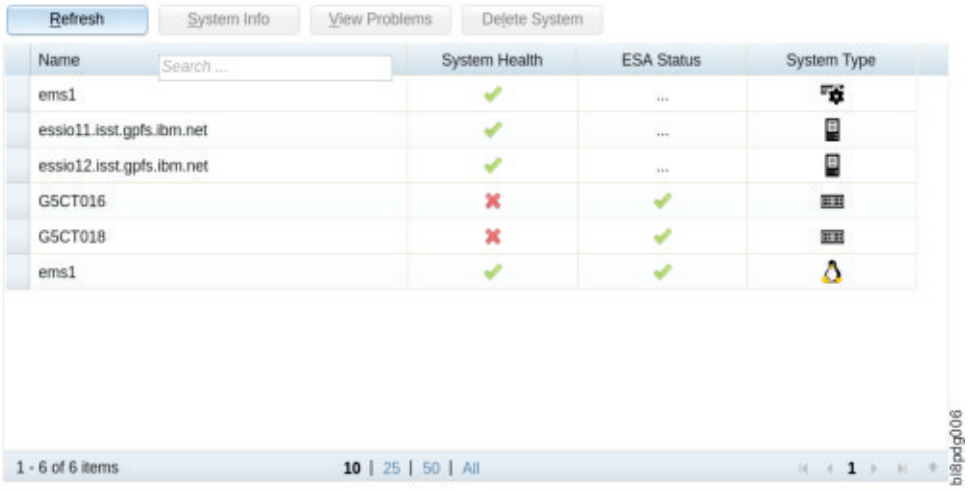


Figure 6. ESA portal showing enclosures with drive replacement events

Another example is POWER9 or POWER8 hardware failure which could be an ESS 5000 or ESS Legacy IO node or EMS node or protocol node hardware failure. Hardware failure event could be DIMM failure, power supply failure, or fan failure in POWER9 or POWER8 nodes. Any of the failure events reported by the POWER nodes are recorded in OPAL log and ESS hardware call home function reads the OPAL events which need service and it is a call home event.

ESS 5000, ESS 3000 and ESS LegacyHardware failure event raised in OPAL log reported by event type:

- `nodeEvent` - Any POWER9 node hardware failure which is reported in the OPAL log is a call home event that is reported with this event.

Similarly, ESS 3000 x86 nodes can also report any of the hardware failure events with **`mmhealth`**. The ESS 3000 call home depends on software call home and **`mmhealth`**. In other words, software call home must be configured and **`mmhealth`** must detect the issue with hardware to get it reported to ESA and a PMR getting opened.

While running **`esscallhomeconf`** command, make sure to not use the **`--no-swcallhome`** switch in the ESS 3000 environment. Otherwise, the ESS 3000 hardware failure events are not reported to ESA and a PMR does not get opened.

ESS 3000 supported hardware events (Need software call home to be configured):

- `bootDrvFail` - When boot drive failed at canister.
- `canFailed` - When canister failed.
- `bootDrvMissing` - When boot drive missing.
- `bootDrvSmtFailed` - When boot drive smart failed.
- `canFanFailed` - When canister FAN stopped working.
- `fanFailed` - When FAN failed.
- `psFailed` - Power supply failure.
- `psFanFailed` - Power supply FAN failure.

Problem details section of ESA

The problem descriptions of the events can be seen by selecting the endpoint. You can select an endpoint by clicking the red X.

The following figure shows an example of the problem description.

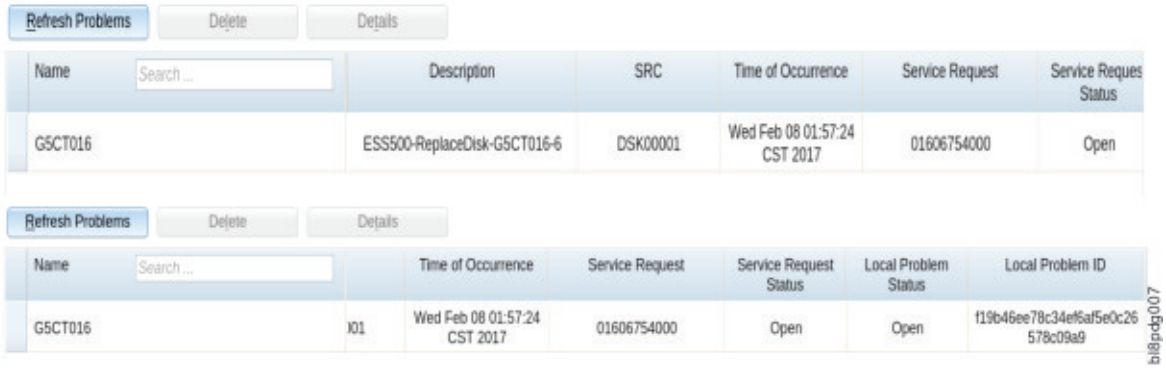


Figure 7. Problem Description

Name

It is the serial number of the enclosure containing the drive to be replaced.

Description

It is a short description of the problem. It shows ESS version or generation, service task name and location code. This field is used in the synopsis of the problem (PMR) report.

SRC

It is the Service Reference Code (SRC). An SRC identifies the system component area. For example, DSK XXXXX, that detected the error and additional codes describing the error condition. It is used by the support team to perform further problem analysis, and determine service tasks associated with the error code and event.

Time of Occurrence

It is the time when the event is reported to the ESA. The time is reported by the endpoints in the UTC time format, which ESA displays in local format.

Service request

It identifies the problem number (PMR number).

Service Request Status

It indicates reporting status of the problem. The status can be one of the following:

Open

No action is taken on the problem.

Pending

The system is in the process of reporting to the IBM support.

Failed

All attempts to report the problem information to the IBM support has failed. The ESA automatically retries several times to report the problem. The number of retries can be configured. Once failed, no further attempts are made.

Reported

The problem is successfully reported to the IBM support.

Closed

The problem is processed and closed.

Local Problem ID

It is the unique identification or event id that identifies a problem.

Problem Details

Further details of a problem can be obtained by clicking the **Details** button. The following figure shows an example of a problem detail.

Problem Summary

Property	Value
Description	ESS500-ReplaceDisk-G5CT018-5
Error Code	DSK00001
Local Problem Status	Open
Problem ID	53c76032dbb54069a28db04a7c229bc3
Is Test Problem?	false
Problem Occurence Date/Time	2/8/17 1:57 AM

Transmission Summary

Property	Value
Service Information Sent to IBM support	Yes
Last Attempt to Send	2/8/17 1:57 AM
Number of Attempts	1

Service request information

Property	Value
Problem Severity	
Service Request Number	01605754000
Service Request Status	Open
Last Changed	2/8/17 1:57 AM

bl8pdg008

Figure 8. Example of a problem summary

If an event is successfully reported to the ESA, and an event ID is received from the ESA, the node reporting the event uploads additional support data to the ESA that are attached to the problem (PMR) for further analysis by the IBM support team.

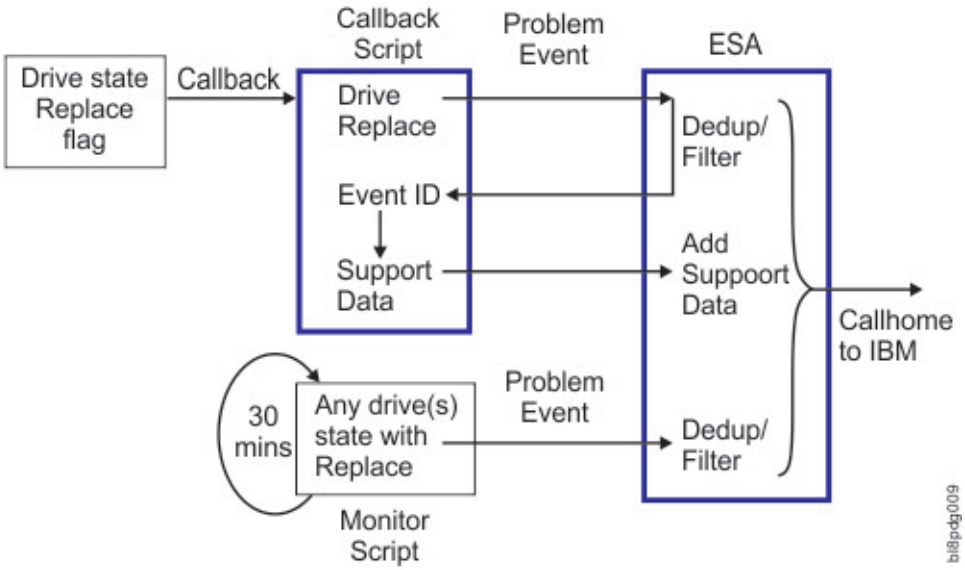


Figure 9. Call home event flow

The callback script logs information in the `/var/log/messages` file during the problem reporting episode. The following examples display the messages logged in the `/var/log/message` file generated by the `essio11` node:

- Callback script is invoked when the drive state changes to `replace`. The callback script sends an event to the ESA:

```
Feb 8 01:57:24 essio11 esscallhomeevent: [I] Event successfully sent
for end point G5CT016, system.id 524e48d68ad875ffbeeec5f3c07e1acf,
 location G5CT016-6, fru 00LY195.
```

- The ESA responds by returning a unique event ID for the system ID in the json format.

```
Feb 8 01:57:24 essio11 esscallhomeevent:
{#012 "status-details": "Received and ESA is processing",
#012 "event.id": "f19b46ee78c34ef6af5e0c26578c09a9",
#012 "system.id": "524e48d68ad875ffbeeec5f3c07e1acf",
#012 "last-activity": "Received and ESA is processing"
#012}
```

  Note: Here #012 represents the new line feed \n.

- The callback script runs the **ionodedatacol.sh** script to collect the support data. It collects the last 10000 lines of `mmfs.log.latest` file and the last 24 hours of the kernel messages in the journal into a .tgz file.

```
Feb 8 01:58:15 essio11 esscallhomeevent: [I] Callhome data collector
/opt/ibm/gss/tools/samples/ionodechdatacol.sh finished

Feb 8 01:58:15 essio11 esscallhomeevent: [I] Data upload successful
for end point 524e48d68ad875ffbeeec5f3c07e1acf
and event.id f19b46ee78c34ef6af5e0c26578c09a9
```

Call home monitoring of ESS 5000, ESS 3000 and ESS Legacy systems and their disk enclosures

A callback is a one-time event. Therefore, it is triggered when the disk state changes to `replace`. If ESA misses the event, for example if the EMS node is down for maintenance, the call home event is not generated by ESA.

To mitigate this situation, the `callhomemon.sh` script is provided in the `/opt/ibm/gss/tools/samples` directory of the EMS node. This script checks for pdisks that are in the `replace` state, and sends an event to ESA to generate a call home event if there is no open PMR for the corresponding physical drive. This script can be run on a periodic interval. For example, every 30 minutes.

In the EMS node, create a cronjob as follows:

1. Open crontab editor by using the following command:

```
# crontab -e
```

2. Setup a periodic cronjob by adding the following line:

```
*/30 * * * * /opt/ibm/gss/tools/samples/callhomemon.sh
```

3. View the cronjob by using the following command:

```
# crontab -l
*/30 * * * * /opt/ibm/gss/tools/samples/callhomemon.sh
```

The call home monitoring protects against missing a call home due to ESA missing a callback event. If a problem report is not already created, the call home monitoring ensures that a problem report is created.

Note: When the call home problem report is generated by the monitoring script, as opposed to being triggered by the callback, the problem support data is not automatically uploaded. In this scenario, the IBM support can request support data from the customer.

Note: A PMR is created because of the periodic checking of the replaced drive state. For example, when the callback event is missed, additional support data is not provided for the ESA agent.

Attention:

- In case of ESS 5000 or ESS Legacy systems, if the hardware event reported by OPAL or any disk enclosures disk error event attached to POWER IO nodes, which was missed because ESA was down or due to some other issue in the EMS node, then events can be triggered manually by invoking the **/opt/ibm/gss/tools/samples/callhomemon.sh** script. The **callhomemon.sh** script reports any missed or new hardware event reported by any POWER9 or POWER8 node, which might be a part of ESS 5000 or ESS Legacy cluster (such as POWER EMS Node, POWER protocol nodes, etc.) and ESS 3000 disk enclosures disk failure event only, if it is a part of ESS cluster.
- In case of ESS 3000 systems, if the hardware event reported by **mmhealth** was missed because ESA was down or due to some other issue in the EMS node then event can be re-triggered by using **mmhealth node eventlog --clear** followed by **mmsysmoncontrol restart**. In case the disk enclosures disk event was missed on the ESS 3000 system because the ESA was down or due to some other issue at EMS node then event can be triggered manually by invoking the **/opt/ibm/gss/tools/samples/callhomemon.sh** script. The **callhomemon.sh** script in case of ESS 3000 only re-sends the missed or old disk enclosures disk error event and any missed or new hardware event reported by any POWER9 which may be a part of an ESS 3000 cluster (such as POWER EMS Node, POWER protocol nodes, etc.).

Upload data

The following support data is uploaded when the ESS system disks in enclosures display a drive replace notification on an ESS 5000, ESS 3000, or ESS Legacy system.

- The output of **mmlspdisk** command for the pdisk that is in replace state.
- Additional support data is provided only when the event is initiated as a response to a callback. The following information is supplied in a .tgz file as additional support data:
  - Last 10000 lines of `mmfs.log.latest` from the node which generates the event.
  - Last 24 hours of the kernel messages (from journal) from the node which generates the event.

The following support data is uploaded when the system displays any hardware issue in an ESS 5000 or an ESS Legacy system.

- The output of the **opal_elog_parse** command for the serviceable event that caused failure.
- Additional support data is provided only when the event is initiated as a response to a callback. The following information is supplied in a .tgz file as additional support data:
  - Last 10000 lines of `mmfs.log.latest` from the node which generates the event.
  - Last 24 hours of the kernel messages (from journal) from the node which generates the event.

The following support data is uploaded when the system displays any hardware issue in an ESS 3000 system.

- The output of the **mmhealth** command and the actual component that caused failure.
- Additional support data is provided only when the event is initiated as a response to a callback. The following information is supplied in a .tgz file as additional support data:
  - Last 10000 lines of `mmfs.log.latest` from the node which generates the event.
  - Last 24 hours of the kernel messages (from journal) from the node which generates the event.

Uninstalling and reinstalling the IBM Electronic Service Agent

The ESA agent is preinstalled in the EMS node from the factory.

Issue the following command to remove the rpm if needed:

```
yum remove esagent.pLinux-4.5.5-1.noarch
```

Issue the following command to reinstall the rpm files for the ESA agent:

```
yum localinstall  path/esagent.pLinux-4.5.5-1.noarch.rpm
```

Where the path is `/install/ess/otherpkgs/rhels8/ppc64le/ess`. The path can also be `/opt/ibm/ess/mnt/installer/otherpkgs/rhels7/ppc64le/ess` or `/opt/ibm/ess/mnt/installer/otherpkgs/rhels8/x86_64/ess` if **essmgr** is run to start the container.

Attention:

- The ESA agent requires the Open JDK 8 files to be installed on the provided node by the using the standard RHEL DVD. The `gpfs.java` package must be installed on the node but make sure that `JAVA_HOME` is not pointing to the `gpfs.java` installed directory because `gpfs.java` contains JDK 11 which is not compatible with ESA.
- Make sure that there is no other JDK such as JDK 11 or any other version of the JDK other than JDK 8 (provided by standard RHEL DVD) installed on the EMS node. ESA uses the default Java version which is in default the PATH and it must be Open JDK v8.x. If a JDK version other than JDK 8 is installed and the default PATH to Java uses a different version of JDK, the ESA activation will fail and call home function will not work.
- Administrator can clean any other or older version of JDK by using the **yum remove <install_java>** command and then mount the DVD ISO provided by the ESS binary and get the Open JDK 8 installed on EMS node.
- Do not change the `gpfs.java` package either by re-installing or uninstalling or else GPFS GUI functionality will not work.

Test call home

The configuration and setup for call home must be tested to ensure that the disk replace event can trigger a call home.

The test is composed of three steps:

- ESA connectivity to IBM - Check connectivity from ESA to IBM network. This might not be required if done during the activation.

  ```
  /opt/ibm/esa/bin/verifyConnectivity -t
  ```

- ESA test call home - Test call home from the ESA portal. Go to **All systems** > **System Health** for the endpoint from which you would like to generate a test call home. Click **Send Test Problem**' from the newly opened **Problems** tab.
- ESS call home script setup to ensure that the callback script is set up correctly.

  Verify that the periodic monitoring is set up.

  ```
  crontab -l
  [root@ems1 deploy]# crontab -l

  */30 * * * * /opt/ibm/ess/tools/samples/callhomemon.sh
  ```

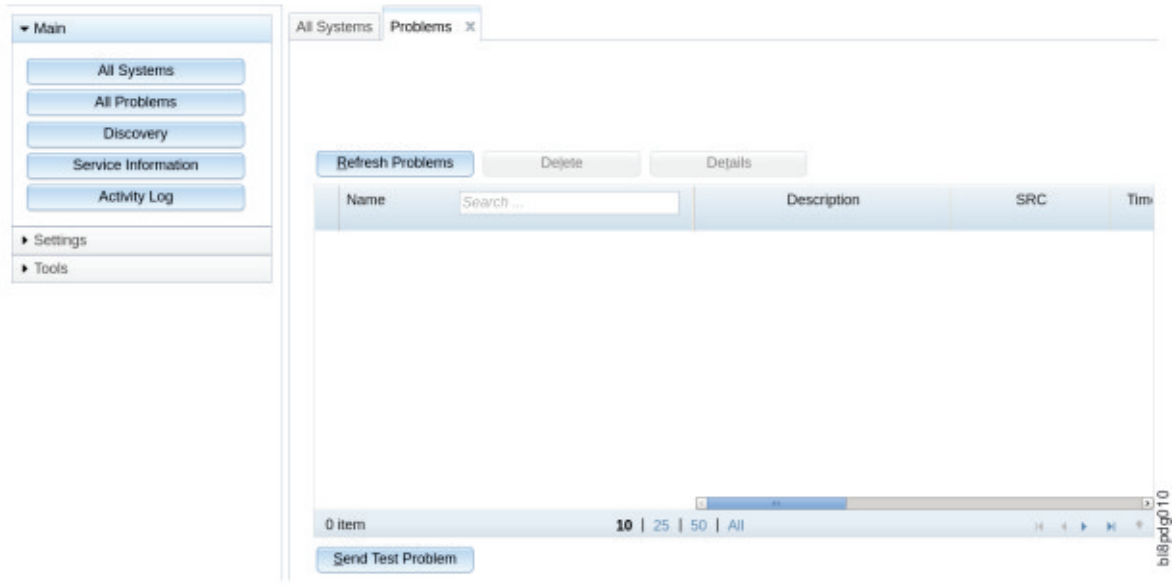


Figure 10. Sending a Test Problem

Callback script test

Verify that the system is healthy by issuing the **gnrhealthcheck** command. You must also verify that the active recovery group (RG) server is the primary recovery group server for all recovery groups. For more recovery group details, see the *IBM Spectrum Scale RAID: Administration* guide.

Example:

To test the callback script, select a pdisk from each enclosure alternating recovery groups. The purpose of the test call home events is to ensure that all the attached enclosures can generate call home events by using both the I/O servers in the building block.

For an ESS 5000 building block, select a pdisk from each enclosure alternating recovery groups. The purpose of the test call home events is to ensure that all the attached enclosures can generate call home events by using both the I/O server nodes in the building block.

In an ESS 5000 system, select a pdisk from each enclosure alternating the paired recovery groups. The purpose of the test call home is to ensure that all the attached enclosures can generate call home events by using both the I/O Server nodes in the building block.

For example, in a SC2 system with 5147-106 enclosures, one can select pdisks e1s001 (left RG, rgL) and e2s106 (right RG, rgR). Similarly, for a SL2 system with 5147-092 enclosures, one can select pdisks e1s02 (left RG, rgL) and e2s92 (right RG, rgR). You must find the corresponding recovery group and active server for these pdisks. Send a disk event to the ESA from the active recovery group server as shown in the following steps:

1. ssh to essio11.

   Here the paired recovery groups are rg1L and rg1R, and the corresponding active I/O server nodes are essio11-ib and essio12-ib.

   Note: Ensure that you state `Test symptom generated by Electronic Service Agent` in the `--eventName` option. Check in the ESA that the enclosure system health is showing the event. You might have to refresh the screen to make the event visible.

   Select the event to see the details:

```
esscallhomeevent --event pdReplacePdisk
--eventName "Test symptom generated by Electronic Service Agent"
--rgName rg1L --pdName e1s001
```

2. ssh to essio12 and run the following command:

```
esscallhomeevent --event pdReplacePdisk
--eventName "Test symptom generated by Electronic Service Agent"
--rgName rg1R --pdName e2s106
```

Figure 11. List of events

Post setup activities

Perform the following post setup activity.

- Delete any test problems.

essinstallcheck enhancement of software and hardware call home

essinstallcheck is now capable of checking and verifying that the systems are configured with software and hardware call home. If the cluster is not yet created, the command skips checking for software call home.

```
# essinstallcheck -N localhost
Start of install check
nodelist:  localhost
Getting package information.
[WARN] Package check cannot be performed other than on EMS node. Checking nodes.
================= Summary of node: localhost ============================
[INFO] Getting system firmware level. May take a long time...
[INFO] Getting system profile setting.
[INFO] Spectrum Scale RAID is not active, cannot get gpfs Installed
version:
[OK] Linux kernel installed:                    3.10.0-1160.11.1.el7.ppc64le
[ERROR] Systemd not at min recommended level: 219-78.el7_9.2.ppc64le
[ERROR] Networkmgr not at min recommended level: 1.18.8-2.el7_9.ppc64le
[OK] Mellanox OFED level:                               MLNX_OFED_LINUX-4.9-2.2.5.1
[OK] IPR SAS FW:                                19512B00
[OK] ipraid RAID level:                         10
[ERROR] ipraid RAID Status: found Degraded expected Optimized
[OK] IPR SAS queue depth:                       64
[ERROR] System Firmware :                       found FW860.81 (SV860_215) expected min
FW860.90 (SV860_226)
[OK] System profile setting:                    scale
[OK] System profile verification PASSED.
[INFO] Cluster not yet created skipping rsyslog check
[OK] Host adapter driver:                       34.00.00.00
Performing Spectrum Scale RAID configuration check.
[OK] New disk prep script: /usr/lpp/mmfs/bin/tspreparenewpdiskforuse
[OK] Network adapter MT4099 firmware: 16.27.2008, net adapter count: 3
[OK] Network adapter firmware
[INFO] Storage firmware check is not required as GPFS cluster does not exist.
[OK] Node is not reserving KVM memory.
[OK] IBM Electronic Service Agent (ESA) is activated for Callhome service.
[OK] Software callhome check skipped as cluster not configured.
End of install check
[PASS] essinstallcheck passed successfully
```

You can view two more lines in the **essinstallcheck** output (in bold face) which mention that ESA is activated (ESA activation indicates that the hardware call home is also configured for this ESS) and software call home has been configured for this node. This is a very important check which enables customers to configure hardware and software call home after the cluster creation and the file system creation is done.

Remember: Enable the hardware and the software call home at the end of the ESS system deployment when the file system is active, nodes are ready to serve the file system, and none of the configuration is pending.

Appendix C. Upgrading the POWER9 firmware

The POWER9 firmware must be upgraded manually. If the firmware is not at the latest level, do the following steps.

1. Copy the firmware `img` file from the container to the POWER9 EMS or the POWER9 protocol node.

   ```
   cd /install/ess/otherpkgs/rhels8/ppc64le/firmware/
   sftp EMSNode
   mput 01VL950_072_045.img
   ```

2. Shut down the container.

   ```
   podman stop ContainerName
   ```

3. Upgrade the firmware.

   a) Run the following command.

   ```
   update_flash -v -f 01VL950_072_045.img
   ```

   b) If there are no issues, execute the update.

   ```
   update_flash -f 01VL950_072_045.img
   ```

   The system restarts and the firmware is upgraded. This process might take 30 - 45 minutes.

Note: If you plan to upgrade the POWER8 EMS firmware, you can retrieve the code from the following location inside the container (the image file will be different):

```
/install/ess/otherpkgs/rhels7/ppc64le/firmware/
```

The level after the upgrade will be SV860_226 (FW860.90).

Appendix D. How to set up chronyd (time server)

Note: The following time server setup documentation is for general reference. You can configure the time server as suitable for your environment. In the simplest example, the EMS host is used as the time server and the I/O nodes (or protocol nodes) are used as clients. Customers might want to have all nodes point to an external time server. Use online references for more detailed instructions for setting up Chrony.

Chrony is the preferred method of setting up a time server. NTP is considered deprecated. Chrony uses the NTP protocol.

For the following example steps, it is assumed that the EMS node is the `chronyd` server and there is no public internet synchronization.

- Do the following steps on the EMS node, outside of the container.
  a) Set the time zone and the date locally.
  b) Edit the contents of the `/etc/chrony.conf` file.

  Note: Replace the server and the allow range with the network settings specific to your setup.

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
#server 0.rhel.pool.ntp.org iburst
#server 1.rhel.pool.ntp.org iburst
#server 2.rhel.pool.ntp.org iburst
#server 3.rhel.pool.ntp.org iburst
server 192.168.7.1 prefer iburst

# Record the rate at which the system clock gains/losses time.
driftfile /var/lib/chrony/drift
local stratum 8
manual

# Allow the system clock to be stepped in the first three updates
# if its offset is larger than 1 second.
makestep 1.0 3

# Enable kernel synchronization of the real-time clock (RTC).
rtcsync

# Enable hardware timestamping on all interfaces that support it.
#hwtimestamp *

# Increase the minimum number of selectable sources required to adjust
# the system clock.
#minsources 2

# Allow NTP client access from local network.
#allow 192.168.0.0/16
allow 192.168.7.0/24

# Serve time even if not synchronized to a time source.
#local stratum 10

# Specify file containing keys for NTP authentication.
#keyfile /etc/chrony.keys

# Specify directory for log files.
logdir /var/log/chrony

# Select which information is logged.
#log measurements statistics tracking
```

  c) Save the changes in `/etc/chrony.conf` file.
  d) Restart `chronyd`.

```
systemctl restart chronyd

chronyc makestep
```

```
chronyc ntpdata
timedatectl
```

- Do the following steps on the client nodes (canister nodes or ESS nodes).

  a) Edit the contents of the `/etc/chrony.conf` file.

  Note: Replace the server and the allow range with the network settings specific to your setup.

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
#server 0.rhel.pool.ntp.org iburst
#server 1.rhel.pool.ntp.org iburst
#server 2.rhel.pool.ntp.org iburst
#server 3.rhel.pool.ntp.org iburst
server 192.168.7.1 prefer iburst

# Record the rate at which the system clock gains/losses time.
server master iburst
driftfile /var/lib/chrony/drift
logdir /var/log/chrony
log measurements statistics tracking

# Allow the system clock to be stepped in the first three updates
# if its offset is larger than 1 second.
makestep 1.0 3

# Enable kernel synchronization of the real-time clock (RTC).
rtcsync

# Enable hardware timestamping on all interfaces that support it.
#hwtimestamp *

# Increase the minimum number of selectable sources required to adjust
# the system clock.
#minsources 2

# Allow NTP client access from local network.
#allow 192.168.0.0/16
#allow 192.168.7.0/24

# Serve time even if not synchronized to a time source.
#local stratum 10

# Specify file containing keys for NTP authentication.
#keyfile /etc/chrony.keys

# Specify directory for log files.
logdir /var/log/chrony

# Select which information is logged.
#log measurements statistics tracking
```

  b) Save the changes in the `/etc/chrony.conf` file.

  c) Restart `chronyd`.

```
systemctl restart chronyd

chronyc makestep
chronyc ntpdata
timedatectl
```

Appendix E. ESS protocol node deployment by using the IBM Spectrum Scale installation toolkit

The following guidance is for adding a protocol node after storage deployment in an ESS environment.

Note: The following instructions for protocol node deployment by using the installation toolkit is just an example scenario. For detailed and latest information, see the following topics.

- Installing IBM Spectrum Scale on Linux nodes with the installation toolkit
- Configuring the CES and protocol configuration

Prerequisites

- During file system creation, adequate space is available for CES shared root file system. For more information, see "During file system creation, adequate space is available for CES shared root file system" on page 51
- ESS container has the protocol node management IP addresses defined. For more information, see "ESS container has the protocol node management IP addresses defined" on page 51.
- ESS container has the CES IP addresses defined. For more information, see "ESS container has the CES IP addresses defined" on page 52.

During file system creation, adequate space is available for CES shared root file system

In a default ESS setup, you can use the Ansible based file system task to create the recovery groups, vdisk sets, and file system. By default, during this task, 100% of the available space is attempted to be consumed. If you plan to include protocol nodes in your setup, you must leave enough free space for the required CES shared root file system. Use the **`--size`** flag to adjust the space consumed accordingly.

For example: **`essrun -G ess_ppc64le filesystem --suffix=-hs --size 80%`**

Running this command leaves approximately 20% space available for the CES shared root file system or additional vdisks. If you are in a mixed ESS 3000 and ESS 5000 environment, you might not use the **`essrun filesystem`** task due to more complex storage pool requirements. In that case, when using **`mmvdisk`**, make sure that you leave adequate space for the CES shared root file system. The CES shared root file system requires around 20 GB of space for operation.

ESS container has the protocol node management IP addresses defined

Before running the ESS container make sure to add the protocol node management IP addresses to `/etc/hosts`. These IP addresses are given to the SSR through the TDA process and they are already set. The customer needs to define host names and add the IP addresses to the EMS node `/etc/hosts` file before running the container.

You also need to define the high-speed IP address and host names. The IP addresses get set when running the Ansible network bonding task but the host names and IP addresses must be defined in `/etc/hosts` before the container starts. The high-speed host names must add a suffix of the management names. The IP addresses are user definable. Consult the network administrator for guidance.

For example:

```
# Protocol management IPs
192.168.45.23 prt1.localdomain prt1
192.168.45.24 prt2.localdomain prt2
# Protocol high-speed IPs
11.0.0.4 pr1-hs.localdomain prt1-hs
11.0.0.5 pr2-hs.localdomain prt2-hs
```

Note: `localdomain` is an example domain. The domain must be changed and also match that of the other nodes.

ESS container has the CES IP addresses defined

The final item that must be defined before starting the ESS container are the CES IP addresses. The following example shows the usage of two IP addresses per node over the high-speed network. Consult the IBM Spectrum Scale documentation for best practices.

```
11.0.0.100 prt_ces1.localdomain prt_ces1
11.0.0.101 prt_ces2.localdomain prt_ces2
11.0.0.102 prt_ces3.localdomain prt_ces3
11.0.0.103 prt_ces4.localdomain prt_ces4
```

Starting state in the example scenario

- ESS storage is deployed and configured.
- Adequate space (approximately 20 GB) is available for CES shared root file system.
- Protocol node required host names and IP addresses is defined on the EMS prior to starting the container.
- You are logged in from the ESS container.

Instructions for deploying protocol nodes in an ESS environment

Do the following steps from the ESS container.

1. Ping the management IP addresses of the protocol nodes.

   ```
   ping IPAdress1,...IPAddressN
   ```

   Each protocol node must respond to the ping test indicating they have an IP address set and it is on the same subnet as the container.

2. Run the config load task.

   ```
   essrun -N ems1,essio1,essio2,prt1,prt2 config load -p RootPassword
   ```

   If you have more than one node, you can specify them in a comma-separated list. Make sure that you add all ESS nodes in this **`config load`** command before continuing.

3. Create network bonds.

   Note: Make sure that the nodes are connected to the high-speed switch before doing this step.

   ```
   essrun -N prt1,prt2 network --suffix=-hs
   ```

4. Install the CES shared root file system.

   ```
   essrun -G ess_ppc64le filesystem --suffix=-hs --ces
   ```

5. Install IBM Spectrum Scale by using the installation toolkit and set up CES.

   Use the IBM Spectrum Scale documentation for installing IBM Spectrum Scale by using the installation toolkit and for enabling the required services for the customer environment. For more information, see:

   - Using the installation toolkit to perform installation tasks: Explanations and examples.
   - Adding CES HDFS nodes into the centralized file system.
   - ESS awareness with the installation toolkit.

Appendix F. Sample scenario: ESS 3000 and ESS 5000 mixed cluster and file system

Use these instructions for setting up ESS 3000 and ESS 5000 mixed cluster and file system.

The following high-level tasks need to be done for setting up ESS 3000 and ESS 5000 mixed cluster:

- Deploy an ESS 3000 system (including cluster, file system, GUI).
- Deploy an ESS 5000 system (adding to cluster, create recovery groups, etc.).
- Create the ESS 5000 vdisks and add to the existing file system.
- Create a policy file.
- Adjust sensors.
- Add ESS 5000 nodes to the GUI.

Note: These instructions contain summarized steps and references to documents that cover the items in more detail. The goal is to give an example scenario to help clients understand aspects of this procedure. At the end of this procedure, if you have POWER9 protocol nodes, for guidance in implementing them into your environment, see Appendix E, “ESS protocol node deployment by using the IBM Spectrum Scale installation toolkit,” on page 51.

Prerequisites

- SSR has completed code 20 on both the ESS 3000 and ESS 5000 nodes (including EMS)

  SSR works on Power® nodes and the EMS node first, then the ESS 3000 system.
- Public connection setup on C11-T3 (f3 connection on EMS)
- ESS 3000 and ESS 5000 nodes have been added to `/etc/hosts`
  - Low-speed names FQDNs , short names, and IP addresses
  - High-speed names FQDNs, short names, and IP addresses (add suffix of low-speed names)
- Host name and domain set on EMS
- Latest code for ESS 3000 and ESS 5000 stored in /home/deploy on EMS
- For information on how to deploy the ESS system, see ESS 3000 Quick Deployment Guide.
- For information on using the **mmvdisk** command, see mmvdisk in ESS documentation.

Summarized version of steps for deploying ESS 3000 building blocks

1. Extract the ESS 3000 installation package: **tar zxvf** *ESS3000InstallationPackage*
2. Accept the license and deploy the container: **sh ess3000_6.1.0.1_0509-21_dae.sh --start-container**

After logging in to the container, do the following steps:

1. Run the config load command.

   ```
   essrun -N ESS3000Node1,ESS3000Node2,EMSNode config load -p RootPassword
   ```

   Note: Use the low-speed names.
2. If required, update the EMS node.

   ```
   essrun -N EMSNode update --offline
   ```
3. Update the ESS 3000 nodes.

```
essrun -N ESS3000Node1,ESS3000Node2 update --offline
```

4. Create network bonds.

```
essrun -N ESS3000Node1,ESS3000Node2,EMSNode network --suffix=Suffix
```

5. Create the cluster.

```
essrun -G ESS3000NodeGroup cluster --suffix=Suffix
```

Note: To obtain the group name, use **lsdef -t group**.

6. Add the EMS node to the cluster.

```
essrun -N ESS3000Node1 cluster --suffix=Suffix --add-ems EMSNode
```

7. Create the file system.

```
essrun -G ESS3000NodeGroup filesystem --suffix=Suffix
```

Note: This command creates a combined data and metadata vdisk in the system pool. The file system name must be fs3k.

Type `exit` and press `Enter` to exit the container. Proceed with the instructions on how to setup the collector, sensors, and run the GUI wizard.

The current ESS 3000 container should be in the `stopped` state. To confirm, use the **podman ps -a** command.

If the ESS 3000 container is not in the `stopped` state, use the **podman stop** *ContainerName* command.

Summarized version of steps to add ESS 5000

1. Extract the ESS 5000 installation package: **tar zxvf** *ESS5000InstallationPackage*
2. Verify the integrity of the installation package: **sha256sum -c** *Extractedsha256sumFile*
3. Accept the license and deploy the container: **sh ess5000_6.1.0.1_0510-00_dae.sh --start-container**

After you have logged into the container, do the following steps:

1. Run the config load command.

```
essrun -N ESS5000Node1,ESS5000Node2,ESS3000Node1,ESS3000Node2,EMSNode config load -p ibmesscluster
```

Note: If you plan to add protocol nodes in the cluster, include them in the list of nodes that you are specifying in this command.

2. Update the nodes.

```
essrun -N ESS5000Node1,ESS5000Node2 update --offline
```

3. Create network bonds.

```
essrun -N ESS5000Node1,ESS5000Node2 network --suffix=Suffix
```

4. Add the ESS 5000 nodes to the existing cluster.
   a. SSH to one of the ESS 5000 I/O server nodes. For example:

```
ssh ESS5000Node1
```

   b. Run this command.

```
essaddnode -N ESS5000Node1-hs,ESS5000Node2-hs --cluster-node ESS3000Node-hs --nodetype ess5k --accept-license
```

Note:

- Use the high-speed names.
- If there is an error, you might need to log in to each ESS 5000 node and start GPFS.

```
mmbuildgpl
mmstartup
```

Type `exit` and press `Enter` to exit the container. Running these commands, takes you to the ESS 5000 node.

5. Create **mmvdisk** artifacts.
   a. Create the node class.

```
mmvdisk nc create --node-class ess5k_ppc64le_mmvdisk -N
ListOfESS5000Nodes_highspeedsuffix
```

   b. Configure the node class.

```
mmvdisk server configure --nc ess5k_ppc64le_mmvdisk --recycle one
```

   c. Create recovery groups.

```
mmvdisk rg create --rg ess5k_rg1,ess5k_rg2 --nc ess5k_ppc64le_mmvdisk
```

   d. Define vdiskset.

```
mmvdisk vs define --vs vs_fs5k_1 --rg ess5k_rg1,ess5k_rg2 code 8+2p --bs 16M --ss 80% --
nsd-usage dataOnly --sp data
```

   e. Create vdiskset.

```
mmvdisk vs create --vs vs_fs5k_1
```

   f. Add vdiskset to the file system.

```
mmvdisk fs add --file-system fs3k --vdisk-set vs_fs5k_1
```

   g. Add the policy file.

   Define your policy file. This can be used to move data from the system pool to the data pool when thresholds hit. For more information, see Overview of policies.

   You can also use the GUI to define policies. For more information, see Creating and applying ILM policy by using GUI.

   The following example rule ingests the writes on the ESS 3000 and moves the data to ESS 5000 when it reaches 75% capacity on the ESS 3000:

   - Add callback for automatic movement of data between pools:

```
mmaddcallback MIGRATION --command /usr/lpp/mmfs/bin/mmstartpolicy --event
lowDiskSpace,noDiskSpace --parms "%eventName %fsName"
```

   - Write the policy into a file with the following content:

```
RULE 'clean_system' MIGRATE FROM POOL 'system' THRESHOLD(75,25) WEIGHT(KB_ALLOCATED) TO
POOL 'data'
```

   Note: You need to understand the implications of this rule before applying it in your system. When capacity on ESS 3000 reaches 75%, it migrates files (larger ones first) out of the system pool to the data pool until the capacity reaches 25%.

   h. On the EMS node, run the following command.

```
mmaddcompspec default --replace
```

At this point, add the ESS 5000 nodes to the `pmsensors` list and use the **Edit rack components** option in the GUI to slot the new nodes into the frame.

If you want to add protocol nodes, see Appendix E, “ESS protocol node deployment by using the IBM Spectrum Scale installation toolkit,” on page 51.

Appendix G. Client node tuning recommendations

IBM Spectrum Scale node configuration is optimized for running IBM Spectrum Scale RAID functions.

ESS cluster node configuration is optimized for running IBM Spectrum Scale RAID functions. Protocols, other gateways, or any other non-ESS services must not be run on ESS management server nodes or I/O server nodes. In a cluster with high IO load, avoid using ESS nodes as cluster manager or filesystem manager. For optimal performance the NSD client nodes accessing ESS nodes should be properly configured. ESS ships with `gssClientConfig.sh script` located in `/usr/lpp/mmfs/samples/gss/` directory. This script can be used to configure the client as follows:

```
/usr/lpp/mmfs/samples/gss/gssClientConfig.sh <Comma Separated list of
client nodes or nodeclass>
```

You can run the following to see configuration parameter settings without setting them:

```
/usr/lpp/mmfs/samples/gss/gssClientConfig.sh -D
```

After running this script, restart GPFS on the affected nodes for the optimized configuration settings to take effect.

Important: Do not run **`gssClientConfig.sh`** unless you fully understand the impact of each setting on the customer environment. Make use of the `-D` option to decide if all or some of the settings might be applied. Then, individually update each client node settings as required.

Appendix H. ESS 5000 Capacity upgrade flow

For customers who are looking to add storage to an existing building block, this option is now supported. The goal of a capacity upgrade is to expand a customer's available storage, either by expanding an existing file system or adding new file systems, without the need to buy an entirely new building block. Capacity upgrade is designed to be an online operation and to not interrupt customer workloads.

Supported paths

SL models (5U92)

- SL1 -> SL2
- SL2 -> SL3
- SL2 -> SL4
- SL3 -> SL4
- SL3 -> SL5
- SL4 -> SL5
- SL4 -> SL6
- SL5 -> SL6
- SL5 -> SL7
- SL6 -> SL7

SC models (4U106)

- SC1 -> SC2
- SC2 -> SC3
- SC2 -> SC4
- SC3 -> SC4
- SC3 -> SC5
- SC4 -> SC5
- SC4 -> SC6
- SC5 -> SC6
- SC5 -> SC7
- SC6 -> SC7
- SC6 -> SC8
- SC7 -> SC8
- SC7 -> SC9
- SC8 -> SC9

Prerequisites

1. All new or existing building blocks must be at ESS 6.1.0.0 or later. If there are protocol nodes in the setup, they must also be upgraded to the matching ESS version.
2. If space needs to be made, for example for moving of the EMS, this has to be planned for accordingly.
3. LBS must wear an ESD wrist band when physically working on the hardware (like plugging in SAS cables).

Capacity upgrade considerations

- Do not try to configure call home before the capacity upgrade is complete, that is until resizing is done.
- You can perform additional capacity upgrades while the DA's are rebalancing.
- You can restripe the file system while the DA's are rebalancing.
- Although it is recommended, you do not have to rebalance the file system if NSDs are added during the capacity upgrade.
- For customers who have call home enabled, remember to re-configure call home after the capacity upgrade is complete. This is the last step in the following flow or you can use the instructions in the call home configuration appendix for guidance.

SAS cable plug-in tips

- Unlatch the cable arm from the I/O server node into which you will be plugging in the SAS cable.
- Remove the blue cap.
- Make sure the location code label from the cable matches the port location code and the port number.
- Remove the cap from the SAS cable connector and plug it into the port.
- You should hear a click when the cable is inserted correctly.

SSR tasks

SSR is responsible for the following tasks.

1. Code 20 of the new enclosures - replacing parts as needed.
2. Running or labeling the new SAS cable connections.
3. Potentially making space in the frame - Moving the EMS.

SSR is not responsible for checking system health using **`essutils`** like in a rackless or a rackful solution.

LBS tasks

LBS is responsible for the following tasks.

1. Upgrade of ESS 6.1.0.0 - prior to the capacity upgrade engagement.
2. Post capacity upgrade health checks.
3. Plugging the SAS cables into the adapters and enclosures.
4. Performing capacity upgrade software functions such as conversion and resizing.
5. New storage management functions such as adding new space to existing file system and creating a new file system.
6. Restriping the file system.
7. Replacing any bad parts such as disks or cables.
8. Pre and post engagement operations

Flow

TDA process ensures that the customer is prepared for the capacity upgrade. Considerations such as if there is enough room in the rack or usage of the file system space are planned out.

LBS

1. LBS performs normal ESS software upgrade. Customer must be at ESS 6.1.0.0 for the capacity upgrade. This upgrade is treated as a separate engagement than the future capacity upgrade operation.

=== Capacity upgrade begins ===

SSR

1. The SSR arrives at the customer site. If the EMS needs to be moved, the SSR shuts down GPFS and powers down the server to move. For more information, see *Shutting down and powering up ESS* in *ESS 5.3.x Quick Deployment Guide*.
2. The SSR places the new enclosures in the rack and establishes power connection. Based on the lights, the SSR performs a code 20 operation. If lights indicate any problem, they might need to take a service action.
3. The SSR runs the new SAS cable connections and labels in a bundle and hooks them to the frame. Later when LBS comes, they simply plug in the connections when required in the flow.
4. The SSR places the EMS (if required) back into the existing frame or a new frame. Network connections and power are restored. Once the server is powered on, the SSR (or customer) can start GPFS to return the EMS back into the cluster.

LBS

1. Power on the new enclosure(s).
   - For SLx or SCx, the power cord should be connected. Press the switch to turn on the enclosures.
2. Verify that the system is converted to mmvdisk.
   a. **`mmvdisk nodeclass list`** - This command shows if the **`mmvdisk`** node class exists.

      This command can be run on any node in the cluster.
3. Upon arrival LBS should first perform the normal upgrade-related system verification steps. Run the following from the EMS:
   a. **`gnrhealthcheck`** - This command determines if there are any issues in various areas of ESS. Any problems that show up must be addressed before capacity upgrade starts. Run this command from any node in the cluster. This command only needs to be run once.
   b. **`gssinstallcheck -N localhost`** - This command checks the system to ensure all components match ESS 6.1.0.0 levels. Run from each node in the cluster including protocol nodes.
   c. **`mmhealth node show -N all --verbose`** - This command shows any system health related issues to address. Run this command from any node in the cluster.
   d. Look for any events or tips that are open in the GUI. These also show up when you issue **`mmhealth`** but it is good to check in the GUI as well. The GUI is run from the EMS node.
   e. **`/usr/sbin/opal-elog-parse -s`** - Run this command from each node in the cluster to check for any serviceable events.
4. Verification steps:
   a. **`mmgetstate -a`** - Issue this command to ensure that all daemons are active.
   b. **`mmlsmount all -L`** - Issue this command to ensure that all mount points are still up. The file system must only be mounted on the EMS and protocol nodes (if applicable).

   After these issues are resolved, capacity upgrade can begin.
5. Start by moving both recovery groups to `essio2-hs`.

   Note: The following recovery group names are examples.

   Move the recovery group in the current I/O server node to the peer I/O server node in the same building block.
   a. To list the recovery groups and the current master server, run:

```
mmvdisk recoverygroup list
```

b. To move the recovery group from the current active I/O server node (`rg_essio1-hs`) to the peer I/O server node (`essio2-hs`) in the same building block, run the following commands in the shown order:

```
mmvdisk recoverygroup change --recovery-group rg_essio1-hs --active essio2-hs
```

Running **mmvdisk recoverygroup list** should show both RGs actively managed by `essio2-hs`.

6. Plug in the SAS cables for essio1 on the server and enclosure ends. Shut down GPFS, only on the server just modified, and then reboot the I/O node. Wait for 5 minutes for the node to reboot and paths to be rediscovered. Run the following commands to ensure that essio1 has discovered the new enclosures.

   Note: Before shutting down GPFS, make sure that autoload is turned off (**mmchconfig autoload=no**).

   a. **gssstoragequickcheck -N localhost**

   b. **gssfindmissingdisks -N localhost**

   Both commands should return with no issues and recognize the new enclosure and disk counts. The paths should also be without error. After this is complete, start IBM Spectrum Scale on the node in question by using **mmstartup**. After determining that IBM Spectrum Scale is `active` by using **mmgetstate** proceed to the next step.

7. Move the recovery group ownership to `essio1-hs`. Use the same commands as used in this step but make sure to use the correct node name (`essio1-hs`).

   After the preceding steps are complete, new enclosures have been successfully cabled to both servers, proceed with the following final steps.

8. Rebalance both recovery groups by running from any node in the storage cluster.

   a. **mmvdisk rg list**

   b. **mmvdisk recoverygroup change --recovery-group** *rg1* **--active essio1-hs**

   c. **mmvdisk recoverygroup change --recovery-group** *rg2* **--active essio2-hs**

   d. Check that the ownership has changed using the **mmvdisk recoverygroup list** command.

9. Perform the system verification steps again before proceeding.

10. Update enclosure and drive firmware. If there are any issues, you should stop and replace any disks or enclosures that could not be updated for some reason.

    `CurrentIoServer` implies running the command from either of I/O server nodes in the building block.

    Note: It might take up to an hour for the firmware upgrade to complete. You might notice that the fan starts to run at high speed. This is a known issue.

    a. `CurrentIoServer$` **mmchfirmware --type storage-enclosure**

    b. `CurrentIoServer$` **mmchfirmware --type drive**

    c. **mmhealth node show -N all --verbose**- This command shows any system health related issues to address. (Run from any node in the storage cluster.)

    d. **gnrhealthcheck** - This command determines if there are any issues in various areas of ESS. Any problems that show up must be addressed before capacity upgrade starts.

11. Add new storage into recovery groups.

    mmvdisk rg resize –-rg rg_essio1-hs,rg_essio2-hs -v no

12. Verify that the new storage is available and the DA is rebalancing.

```
mmvdisk recoverygroup list --recovery-group rg1 --all
mmvdisk recoverygroup list --recovery-group rg2 --all
```

Run for both recovery groups. Note that the additional free space available in the DA and the background task for each DA is showing as `'rebalance'`.

On the EMS node, update the component database.

```
mmaddcompspec default --replace
```

13. Start up the GUI and use **Edit rack components** to have the GUI discover the new topologies and make changes to the frame accordingly. Changes such as modify ESS model to consume more U space, move EMS, and so on.
14. Reconfigure call home.

```
esscallhomeconf -E ems1 -N EMSNode,IONode1,IONode2 --suffix=-hs --register=all
```

At this point, discussions with the customers need to occur on what to do with the free space.

1. Add to the existing file system?
   a. See the add building block flow in *ESS 5.3.x Quick Deployment Guide* for tips on creating new NSDs and adding to an existing file system.
   b. Consider file system restripe at the end which might take time. (**mmrestripefs** *FileSystem* **-b**)
2. Create a new file system?
   - See the installation section on how to use **essrun** on creating a new file system from inside the container. You may also use **mmvdisk** commands directly to perform this operation.

Appendix I. Switch VLAN configuration instructions

This topic describes the instructions that are needed to configure an IBM cumulus switch VLAN.

The IBM cumulus switch would be preconfigured from manufacturing with proper VLAN that includes the following:

- Service/FSP/BMC VLAN
  - Blue network - Bottom ports
  - VLAN 101
- Management/Provisioning VLAN
  - Yellow network - Top ports
  - VLAN 102
- IBM Elastic Storage System special ports
  - Ports 1 - 12
  - Trunk ports
  - Default routes traffic to management VLAN
  - Packets with VLAN tag 101 routed to service network.

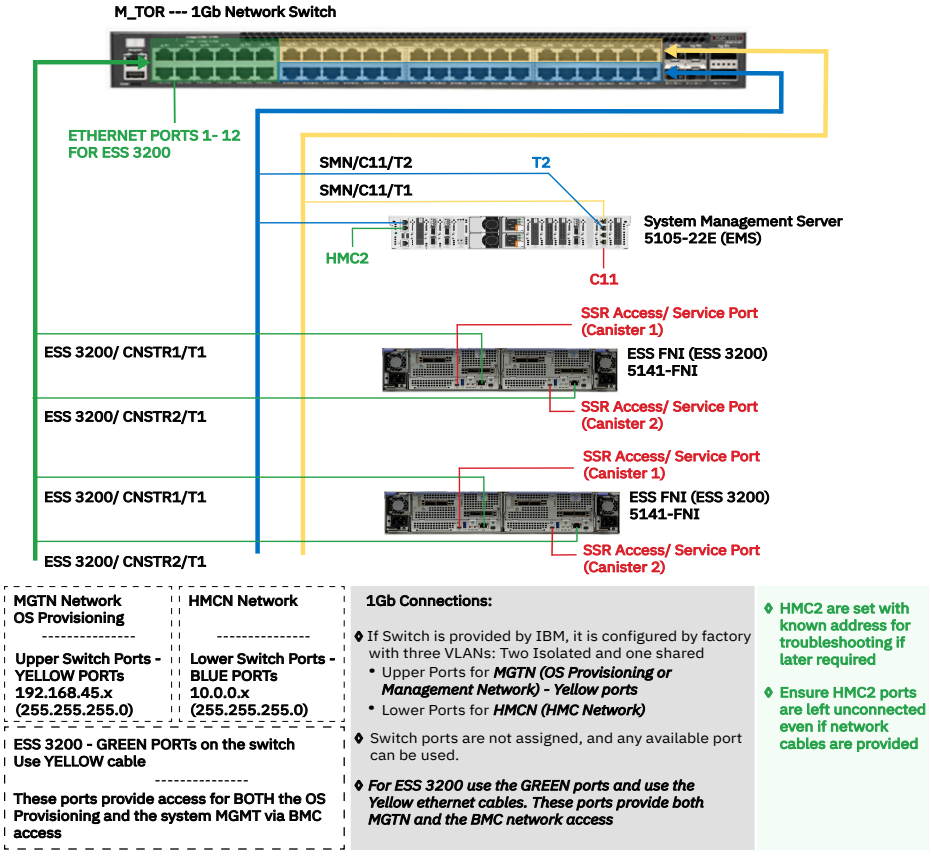


Figure 12. 1 Gb network switch

Procedure to change switch default password

Use the following steps to change switch default password.

1. Verify the 11S label at the back of the switch as shown in the following figure.

Figure 13. 11S label

Note: The required software update is cumulus-3.7.12a.

2. Log in to the switch by using the following default credentials and press the Enter key.
   - User ID: Cumulus
   - Password: CumulusLinux!
3. Use the following command to display the 11S serial number.

```
cumulus@1Gsw:~$ decode-syseeprom | grep Serial | awk '{print $5}'| cut --complement -c -3
```

   The system displays the 11S serial number similar to the following:

```
01FT690YA50YD7BGABX
```

4. Change the default password to the 11S password by using the following command:

```
cumulus@accton-1gb-mgmt:~$ passwd
```

```
current UNIX password: CumulusLinux!
Enter new UNIX password: <<<Copy and paste the output provided in the 11S serial number
display step.
Retype new UNIX password: <<<Copy and paste the output provided in the 11S serial number
display step.
passwd: password updated successfully.
```

5. Log in through SSH or console and log in with the new 11S password to validate the changes.

   Note: The default password must be set to the 11S serial number **01FT690YA50YD7BGABX**. If not, the password must be **CumulusLinux!**.

Connect the PC to the switch console port

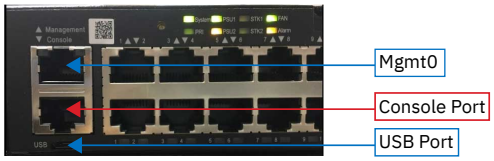


Figure 14. Switch port and switch markings

Connect the PC to the switch console port as follows:

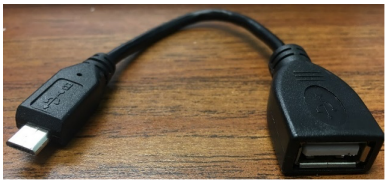
- Connect to the switch by using RJ45 to serial cable.

Figure 15. RJ45 to serial cable and USB to serial cable

- Connect the serial end to the serial to USB cable.

- Connect the USB cable to the PC.

Figure 16. USB cable



Configure the host PC

Configure the host PC as follows:

1. Ensure that the driver for USB to serial cable is connected on the PC.
2. Open the device manager to verify that the COM port is used by the USB to serial cable.
3. Open `putty.exe` and use the COM port to connect to the switch.
4. Configure PuTTY to use as follows:
   a. Baud rate - 115200
   b. Parity - none
   c. Stop bits - 1
   d. Data bits - 8
   e. Flow control - none
5. Power on the switch and wait for the login prompt to show up.
6. Log in by using the following default credentials and press the Enter key.
   - User ID: Cumulus
   - Password: <11S serial number>

   Note: If the switch has default Mellanox user ID and password, then proceed as follows:
   - User ID: Cumulus
   - Password: CumulusLinux!
7. Download the VLAN configuration file by using the following `H48712_interfaces.rtf`:

```
# This file describes the network interfaces available on your system
# and how to activate them. For more information, see interfaces(5).
source /etc/network/interfaces.d/*.intf
# The loopback network interface
auto lo
iface lo inet loopback
# The primary network interface
auto eth0
iface eth0 inet dhcp
# EVEN Ports/Lower ports PVID 101 for FSP network
auto swp14
iface swp14
bridge-access 101
auto swp16
iface swp16
bridge-access 101
auto swp18
iface swp18
bridge-access 101
auto swp20
iface swp20
bridge-access 101
auto swp22
iface swp22
bridge-access 101
auto swp24
iface swp24
bridge-access 101
auto swp26
iface swp26
```

```
bridge-access 101
auto swp28
iface swp28
bridge-access 101
auto swp30
iface swp30
bridge-access 101
auto swp32
iface swp32
bridge-access 101
auto swp34
iface swp34
bridge-access 101
auto swp36
iface swp36
bridge-access 101
auto swp38
iface swp38
bridge-access 101
auto swp40
iface swp40
bridge-access 101
auto swp42
iface swp42
bridge-access 101
auto swp44
iface swp44
bridge-access 101
auto swp46
iface swp46
bridge-access 101
auto swp48
iface swp48
bridge-access 101


# ODD Ports/Upper ports PVID 102 for xCAT network
auto swp13
iface swp13
bridge-access 102
auto swp15
iface swp15
bridge-access 102
auto swp17
iface swp17
bridge-access 102
auto swp19
iface swp19
bridge-access 102
auto swp21
iface swp21
bridge-access 102
auto swp23
iface swp23
bridge-access 102
auto swp25
iface swp25
bridge-access 102
auto swp27
iface swp27
bridge-access 102
auto swp29
iface swp29
bridge-access 102
auto swp31
iface swp31
bridge-access 102
auto swp33
iface swp33
bridge-access 102
auto swp35
iface swp35
bridge-access 102
auto swp37
iface swp37
bridge-access 102
auto swp39
iface swp39
bridge-access 102
auto swp41
iface swp41
bridge-access 102
```

```
auto swp43
iface swp43
bridge-access 102
auto swp45
iface swp45
bridge-access 102
auto swp47
iface swp47
bridge-access 102


# ESS 3200 ports (1 to 12) FSP + OS on single physical port
auto swp1
iface swp1
bridge-pvid 102
bridge-vids 101
auto swp2
iface swp2
bridge-pvid 102
bridge-vids 101
auto swp3
iface swp3
bridge-pvid 102
bridge-vids 101
auto swp4
iface swp4
bridge-pvid 102
bridge-vids 101
auto swp5
iface swp5
bridge-pvid 102
bridge-vids 101
auto swp6
iface swp6
bridge-pvid 102
bridge-vids 101
auto swp7
iface swp7
bridge-pvid 102
bridge-vids 101
auto swp8
iface swp8
bridge-pvid 102
bridge-vids 101
auto swp9
iface swp9
bridge-pvid 102
bridge-vids 101
auto swp10
iface swp10
bridge-pvid 102
bridge-vids 101
auto swp11
iface swp11
bridge-pvid 102
bridge-vids 101
auto swp12
iface swp12
bridge-pvid 102
bridge-vids 101


# Bridge setup
auto bridge
iface bridge
bridge-vlan-aware yes
bridge-ports glob swp1-48
bridge-pvid 101
bridge-pvid 102
bridge-stp off
```

8. Gain sudo rights by using the following command:

```
sudo su -
```

9. Copy the contents of the interface file to the file name `/etc/network/interfaces` and save the file.
10. Reload the interfaces by using the following command:

```
root@cumulus:/etc/network# ifreload -a
root@cumulus:/etc/network# ifquery-a
```

11. Check VLAN setup.

```
net show interface all
```

12. If required, set switch network. It is recommended to set a static IP to log remotely on the switch. For example, 192.168.44.0/24 network IP switch 192.168.44.20, gateway 192.168.44.1.
    - net add interface eth0 IP address 192.168.44.20/24
    - net add interface eth0 IP gateway 192.168.44.1
    - net pending
    - net commit
13. Set the VLAN tag on each server canister. If this document is used, the tag must be 101.

```
# Set tag
/bin/ipmitool lan set 1 vlan id 101
# Confirm tag
/bin/ipmitool lan print 1 | grep -i 'VLAN ID'
```

Non-IBM Cumulus switches

If you have a non-IBM Cumulus switch, use the information above as a general reference on how to modify the switch. The key is to have a designated IBM Elastic Storage System trunk ports that are apart of both VLANs.

Modifying existing Cumulus switches

If you are converting a switch that has already non-ESS 3200 using the switch on any port in the range 1 - 12, you need to evacuate one by one those ports. If you are not using ports in the range 1 - 12, you need to apply the above process.

That means to move the cables on the upper ports in the range 1 - 12 to any free upper port that is not in the range ports 1 - 12. Equally any lower cable plugged to any port in the range 1 - 12 needs to be moved to any lower port not in the range of ports 1 - 12.

You must move one cable at the time and wait until the link LED on the destination port comes up. Once all ports in the range 1-12 are no longer cabled, you can apply the following procedure.

If an existing Cumulus switch must be modified to support IBM Elastic Storage System , the general guidance are as follows:

1. Free up at least two ports (1 IBM Elastic Storage System) on the existing switch. It is better if you can free up a block. Ideally, the current configuration is not scattered where it is easy to convert free ports for IBM Elastic Storage System usage.
2. Take the existing interfaces file from the switch and modify it for the chosen IBM Elastic Storage System ports.
3. Make the modifications to the interfaces file.

```
auto swp10
iface swp10
bridge-pvid 102
bridge-vids 101
```

   Any ports that you designate as IBM Elastic Storage System ports need to have this configuration. Consult the default IBM Elastic Storage System interfaces file for more information.
4. Copy the new interfaces file to the switch.
5. Reload and verify the interfaces.
6. Set the VLAN tags on the IBM Elastic Storage System canisters.

Appendix J. Replacing all POWER8 nodes in an environment with POWER9 nodes in online mode

Replace all POWER8 nodes in your environment with POWER9 nodes in online mode by using this procedure. Cluster and file system remain up during this procedure.

Goal

The goal is to enable the customer or SLS to swap out all the POWER8 ESS nodes in the cluster with the new POWER9 (5105-22E) nodes without taking the cluster or the file system down.

Prerequisites and assumptions

- Existing POWER8 based ESS environment (GLxC – 4U106)
- Existing POWER8 EMS
- All POWER8 nodes (including EMS) will be swapped with POWER9 nodes.
- Existing cluster (at least storage nodes) must be fully updated to ESS 5.3.6.1 levels to match incoming POWER9 nodes.
- There are multiple building-blocks to properly drain BSDs, and maintain quorum, while keeping the cluster active and file system intact.
- There is enough space on the surviving NSDs to store existing data until new NSDs can be added to extend.
- Existing, supported ESS network and client infrastructure:
  - Management switch VLAN'd to spec
  - High-speed switch (Ethernet, InfiniBand, or both)
  - Client nodes with platforms supported by IBM Spectrum Scale
  - Optional POWER8 (5147-22L) protocol nodes in the same cluster as the ESS storage (2 or more protocol nodes)
  - Optional utility nodes, quorum servers, backup nodes, and so on.
- Supported ESS 5000 (POWER9 5105-22E) nodes that match 1-to-1 with the existing POWER8 nodes in the environment that will be swapped out.
- All POWER8 nodes must have the same high-speed adapter counts and port configuration as the POWER8 nodes that will be removed.
- The POWER9 servers should be able to be racked or powered without any changes to the existing infrastructure.

Examples target code levels:

- ESS 5.3.6.1 – For Legacy POWER8 environment
- ESS 6.0.1.1 – For New POWER9 environment

High-level flow

Example environment:

- 1 x POWER8 EMS
- 2 x POWER8 ESS GLxC building-blocks
  - Each building block is in its own failure group
  - Metadata replication between failure groups

- High-speed switch (IB or Ethernet)
- Remote client cluster
- POWER9 EMS (5105-22E) in a box
- 4 x POWER9 IO nodes (5105-22E) in a box

  (All P9components all have required power cables/brackets etc)
- EMS and all 4 IO nodes are quorum nodes
  - EMS is cluster and file system manager
  - Recovery groups are balanced initially per building blocks

Flow:

- Encourage customer or SLS to backup data offline as a precaution. Good practice if space to do it offcluster.
- Take a note of all IP addresses per server. LBS or SSR will need to restore when the POWER9 servers are in place. A new IP address and hostname will be needed for the POWER9 EMS

1. Ensure that the existing POWER8 environment is fully upgraded to ESS 5.3.6.1. Use the ESS 5.3.6.1 Quick Deployment Guide to perform this task.
2. **[Optional]** Upgrade client nodes to matching IBM Spectrum Scale level (5.0.5.3) and OFED (4.9x), if applicable. Move storage release=LATEST and file system format to FULL.
3. Exchange SSH keys between ESS and ESS 5000 (No xCAT running on POWER9 ESS 5000).

   Note: Shut down GUI services on the POWER8 EMS
4. Add the ESS 5000 system to existing ESS cluster.
   - Configure the recovery groups.
   - Configure the vdisk sets.
   - Configure the server lists.
5. Add ESS 5000 disks to the existing ESS file system but do not restripe.
6. Invoke **mmchdisk suspend** on all the disks coming from the ESS so that data will only be written to ESS 5000 disks.
7. Ensure that all GPFS nodes are active, quorum is maintained, NSDs are online, and file system is available. Ensure all storage hardware is healthy before proceeding.
8. Suspend and empty the NSDs from the first building block. Ensure that the data has been properly restriped to the surviving NSDs before continuing.
9. Delete the NSDs, vdisks, recovery groups, and cluster membership of drained NSD nodes (building block).
10. Verify that the old disks are empty with the **mmlsdisk** command.
11. Invoke **mmdeldisk** when all disks are emptied from ESS.
12. **[SSR]** Power off the servers and storage enclosures of the drained building block.
13. **[SSR]** Disconnect power and networking from the building block.
14. **[SSR]** Swap out the building block front the frame

    a. Insert the first pair of POWER9 nodes in their place.

    b. Cable the swapped POWER9 nodes to the ESS 5000 Spec (SAS cables, Ethernet, high speed, and so on).

    Note: POWER9 servers should go through the CSC process.
15. **[SSR]** Power on the new POWER9 building block and connected storage enclosures.
16. **[SSR]** Walk through the ESS 5000 Hardware Guide code 20 flow and set the original IP addresses per node.

17. **[SSR]** Install the POWER9 EMS and walk through the full code 20 flow. Set the new IP addresses provided by the customer.
18. **[Customer or SLS]** Log in to the POWER99 EMS and walk through the *ESS 5000 Quick Deployment Guide* to fully deploy the replaced building block.
19. **[Customer or SLS]** Create network bonds with same IP addresses removed from the POWER8 nodes and new IP address for the EMS itself.
20. **[Customer or SLS]** Add EMS and POWER9 nodes to the existing POWER8 cluster.
21. **[Customer or SLS]** Create NSDs from the POWER9 connected storage.
22. **[Customer or SLS]** Add NSDs to the existing POWER8 file system.
23. **[Customer or SLS]** Force restripe or rebalance.

Repeat the preceding steps for the other POWER8 building blocks besides the POWER9 EMS rack and code 20 portion. When completed, all POWER8 nodes are swapped with POWER9 and a new set of NSDs in place; without losing file system or cluster access.

At the end of the procedure, the EMS node can be removed from the cluster and all POWER8 nodes can be repurposed, perhaps as client nodes.

The GUI, call home, and performance monitoring must be reconfigured to use the new EMS and cluster nodes, although IP addresses of the storage are the same. It is recommended to reconfigure all of these components afresh.

Accessibility features for the system

Accessibility features help users who have a disability, such as restricted mobility or limited vision, to use information technology products successfully.

Accessibility features

The following list includes the major accessibility features in IBM Spectrum Scale RAID:

- Keyboard-only operation
- Interfaces that are commonly used by screen readers
- Keys that are discernible by touch but do not activate just by touching them
- Industry-standard devices for ports and connectors
- The attachment of alternative input and output devices

IBM Knowledge Center, and its related publications, are accessibility-enabled. The accessibility features are described in IBM Knowledge Center (www.ibm.com/support/knowledgecenter).

Keyboard navigation

This product uses standard Microsoft Windows navigation keys.

IBM and accessibility

See the IBM Human Ability and Accessibility Center (www.ibm.com/able) for more information about the commitment that IBM has to accessibility.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing IBM Corporation North Castle Drive Armonk, NY 10504-1785 U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing Legal and Intellectual Property Law IBM Japan Ltd. 19-21,

Nihonbashi-Hakozakicho, Chuo-ku Tokyo 103-8510, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Dept. 30ZA/Building 707
Mail Station P300
2455 South Road,
Poughkeepsie, NY 12601-5400
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment or a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Intel is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

The registered trademark Linux is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Microsoft, Windows, and Windows NT are trademarks of Microsoft Corporation in the United States, other countries, or both.

Red Hat and Ansible are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Glossary

This glossary provides terms and definitions for the IBM Elastic Storage System solution.

The following cross-references are used in this glossary:

- *See* refers you from a non-preferred term to the preferred term or from an abbreviation to the spelled-out form.
- *See also* refers you to a related or contrasting term.

For other terms and definitions, see the IBM Terminology website (opens in new window):

http://www.ibm.com/software/globalization/terminology

B

building block
A pair of servers with shared disk enclosures attached.

BOOTP
See *Bootstrap Protocol (BOOTP)*.

Bootstrap Protocol (BOOTP)
A computer networking protocol that is used in IP networks to automatically assign an IP address to network devices from a configuration server.

C

CEC
See *central processor complex (CPC)*.

central electronic complex (CEC)
See *central processor complex (CPC)*.

central processor complex (CPC)
A physical collection of hardware that consists of channels, timers, main storage, and one or more central processors.

cluster
A loosely-coupled collection of independent systems, or *nodes*, organized into a network for the purpose of sharing resources and communicating with each other. See also *GPFS cluster*.

cluster manager
The node that monitors node status using disk leases, detects failures, drives recovery, and selects file system managers. The cluster manager is the node with the lowest node number among the quorum nodes that are operating at a particular time.

compute node
A node with a mounted GPFS file system that is used specifically to run a customer job. ESS disks are not directly visible from and are not managed by this type of node.

CPC
See *central processor complex (CPC)*.

D

DA
See *declustered array (DA)*.

datagram
A basic transfer unit associated with a packet-switched network.

DCM
See *drawer control module (DCM)*.

declustered array (DA)
A disjoint subset of the pdisks in a recovery group.

dependent fileset
A fileset that shares the inode space of an existing independent fileset.

DFM
See *direct FSP management (DFM)*.

DHCP
See *Dynamic Host Configuration Protocol (DHCP)*.

direct FSP management (DFM)
The ability of the xCAT software to communicate directly with the Power Systems server's service processor without the use of the HMC for management.

drawer control module (DCM)
Essentially, a SAS expander on a storage enclosure drawer.

Dynamic Host Configuration Protocol (DHCP)
A standardized network protocol that is used on IP networks to dynamically distribute such network configuration parameters as IP addresses for interfaces and services.

E

Elastic Storage System (ESS)
A high-performance, GPFS NSD solution made up of one or more building blocks. The ESS software runs on ESS nodes - management server nodes and I/O server nodes.

ESS Management Server (EMS)
An xCAT server is required to discover the I/O server nodes (working with the HMC), provision the operating system (OS) on the I/O server nodes, and deploy the ESS software on the management node and I/O server nodes. One management server is required for each ESS system composed of one or more building blocks.

encryption key
A mathematical value that allows components to verify that they are in communication with the expected server. Encryption keys are based on a public or private key pair that is created during the installation process. See also *file encryption key (FEK)*, *master encryption key (MEK)*.

ESS
See *Elastic Storage System (ESS)*.

environmental service module (ESM)
Essentially, a SAS expander that attaches to the storage enclosure drives. In the case of multiple drawers in a storage enclosure, the ESM attaches to drawer control modules.

ESM
See *environmental service module (ESM)*.

Extreme Cluster/Cloud Administration Toolkit (xCAT)
Scalable, open-source cluster management software. The management infrastructure of ESS is deployed by xCAT.

F

failback
Cluster recovery from failover following repair. See also *failover*.

failover
(1) The assumption of file system duties by another node when a node fails. (2) The process of transferring all control of the ESS to a single cluster in the ESS when the other clusters in the ESS fails. See also *cluster*. (3) The routing of all transactions to a second controller when the first controller fails. See also *cluster*.

failure group
A collection of disks that share common access paths or adapter connection, and could all become unavailable through a single hardware failure.

FEK
See *file encryption key (FEK)*.

file encryption key (FEK)
A key used to encrypt sectors of an individual file. See also *encryption key*.

file system
The methods and data structures used to control how data is stored and retrieved.

file system descriptor
A data structure containing key information about a file system. This information includes the disks assigned to the file system (*stripe group*), the current state of the file system, and pointers to key files such as quota files and log files.

file system descriptor quorum
The number of disks needed in order to write the file system descriptor correctly.

file system manager
The provider of services for all the nodes using a single file system. A file system manager processes changes to the state or description of the file system, controls the regions of disks that are allocated to each node, and controls token management and quota management.

fileset
A hierarchical grouping of files managed as a unit for balancing workload across a cluster. See also *dependent fileset*, *independent fileset*.

fileset snapshot
A snapshot of an independent fileset plus all dependent filesets.

flexible service processor (FSP)
Firmware that provides diagnosis, initialization, configuration, runtime error detection, and correction. Connects to the HMC.

FQDN
See *fully-qualified domain name (FQDN)*.

FSP
See *flexible service processor (FSP)*.

fully-qualified domain name (FQDN)
The complete domain name for a specific computer, or host, on the Internet. The FQDN consists of two parts: the hostname and the domain name.

G

GPFS cluster
A cluster of nodes defined as being available for use by GPFS file systems.

GPFS portability layer
The interface module that each installation must build for its specific hardware platform and Linux distribution.

GPFS Storage Server (GSS)
A high-performance, GPFS NSD solution made up of one or more building blocks that runs on System x servers.

GSS
See *GPFS Storage Server (GSS)*.

H

Hardware Management Console (HMC)
Standard interface for configuring and operating partitioned (LPAR) and SMP systems.

HMC
See *Hardware Management Console (HMC)*.

I

IBM Security Key Lifecycle Manager (ISKLM)
For GPFS encryption, the ISKLM is used as an RKM server to store MEKs.

independent fileset
A fileset that has its own inode space.

indirect block
A block that contains pointers to other blocks.

inode
The internal structure that describes the individual files in the file system. There is one inode for each file.

inode space
A collection of inode number ranges reserved for an independent fileset, which enables more efficient per-fileset functions.

Internet Protocol (IP)
The primary communication protocol for relaying datagrams across network boundaries. Its routing function enables internetworking and essentially establishes the Internet.

I/O server node
An ESS node that is attached to the ESS storage enclosures. It is the NSD server for the GPFS cluster.

IP
See *Internet Protocol (IP)*.

IP over InfiniBand (IPoIB)
Provides an IP network emulation layer on top of InfiniBand RDMA networks, which allows existing applications to run over InfiniBand networks unmodified.

IPoIB
See *IP over InfiniBand (IPoIB)*.

ISKLM
See *IBM Security Key Lifecycle Manager (ISKLM)*.

J

JBOD array
The total collection of disks and enclosures over which a recovery group pair is defined.

K

kernel
The part of an operating system that contains programs for such tasks as input/output, management and control of hardware, and the scheduling of user tasks.

L

LACP
See *Link Aggregation Control Protocol (LACP)*.

Link Aggregation Control Protocol (LACP)
Provides a way to control the bundling of several physical ports together to form a single logical channel.

logical partition (LPAR)
A subset of a server's hardware resources virtualized as a separate computer, each with its own operating system. See also *node*.

LPAR
See *logical partition (LPAR)*.

M

management network
A network that is primarily responsible for booting and installing the designated server and compute nodes from the management server.

management server (MS)
An ESS node that hosts the ESS GUI and xCAT and is not connected to storage. It must be part of a GPFS cluster. From a system management perspective, it is the central coordinator of the cluster. It also serves as a client node in an ESS building block.

master encryption key (MEK)
A key that is used to encrypt other keys. See also *encryption key*.

maximum transmission unit (MTU)
The largest packet or frame, specified in octets (eight-bit bytes), that can be sent in a packet- or frame-based network, such as the Internet. The TCP uses the MTU to determine the maximum size of each packet in any transmission.

MEK
See *master encryption key (MEK)*.

metadata
A data structure that contains access information about file data. Such structures include inodes, indirect blocks, and directories. These data structures are not accessible to user applications.

MS
See *management server (MS)*.

MTU
See *maximum transmission unit (MTU)*.

N

Network File System (NFS)
A protocol (developed by Sun Microsystems, Incorporated) that allows any host in a network to gain access to another host or netgroup and their file directories.

Network Shared Disk (NSD)
A component for cluster-wide disk naming and access.

NSD volume ID
A unique 16-digit hexadecimal number that is used to identify and access all NSDs.

node
An individual operating-system image within a cluster. Depending on the way in which the computer system is partitioned, it can contain one or more nodes. In a Power Systems environment, synonymous with *logical partition*.

node descriptor
A definition that indicates how ESS uses a node. Possible functions include: manager node, client node, quorum node, and non-quorum node.

node number
A number that is generated and maintained by ESS as the cluster is created, and as nodes are added to or deleted from the cluster.

node quorum
The minimum number of nodes that must be running in order for the daemon to start.

node quorum with tiebreaker disks
A form of quorum that allows ESS to run with as little as one quorum node available, as long as there is access to a majority of the quorum disks.

non-quorum node
A node in a cluster that is not counted for the purposes of quorum determination.

O

OFED
See *OpenFabrics Enterprise Distribution (OFED)*.

OpenFabrics Enterprise Distribution (OFED)
An open-source software stack includes software drivers, core kernel code, middleware, and user-level interfaces.

P

pdisk
A physical disk.

PortFast
A Cisco network function that can be configured to resolve any problems that could be caused by the amount of time STP takes to transition ports to the Forwarding state.

R

RAID
See *redundant array of independent disks (RAID)*.

RDMA
See *remote direct memory access (RDMA)*.

redundant array of independent disks (RAID)
A collection of two or more disk physical drives that present to the host an image of one or more logical disk drives. In the event of a single physical device failure, the data can be read or regenerated from the other disk drives in the array due to data redundancy.

recovery
The process of restoring access to file system data when a failure has occurred. Recovery can involve reconstructing data or providing alternative routing through a different server.

recovery group (RG)
A collection of disks that is set up by ESS, in which each disk is connected physically to two servers: a primary server and a backup server.

remote direct memory access (RDMA)
A direct memory access from the memory of one computer into that of another without involving either one's operating system. This permits high-throughput, low-latency networking, which is especially useful in massively-parallel computer clusters.

RGD
See *recovery group data (RGD)*.

remote key management server (RKM server)
A server that is used to store master encryption keys.

RG
See *recovery group (RG)*.

recovery group data (RGD)
Data that is associated with a recovery group.

RKM server
See *remote key management server (RKM server)*.

S

SAS
See *Serial Attached SCSI (SAS)*.

secure shell (SSH)
A cryptographic (encrypted) network protocol for initiating text-based shell sessions securely on remote computers.

Serial Attached SCSI (SAS)
A point-to-point serial protocol that moves data to and from such computer storage devices as hard drives and tape drives.

service network
A private network that is dedicated to managing POWER8 servers. Provides Ethernet-based connectivity among the FSP, CPC, HMC, and management server.

SMP
See *symmetric multiprocessing (SMP)*.

Spanning Tree Protocol (STP)
A network protocol that ensures a loop-free topology for any bridged Ethernet local-area network. The basic function of STP is to prevent bridge loops and the broadcast radiation that results from them.

SSH
See *secure shell (SSH)*.

STP
See *Spanning Tree Protocol (STP)*.

symmetric multiprocessing (SMP)
A computer architecture that provides fast performance by making multiple processors available to complete individual processes simultaneously.

T

TCP
See *Transmission Control Protocol (TCP)*.

Transmission Control Protocol (TCP)
A core protocol of the Internet Protocol Suite that provides reliable, ordered, and error-checked delivery of a stream of octets between applications running on hosts communicating over an IP network.

V

VCD
See *vdisk configuration data (VCD)*.

vdisk
A virtual disk.

vdisk configuration data (VCD)
Configuration data that is associated with a virtual disk.

X

xCAT
See *Extreme Cluster/Cloud Administration Toolkit*.

Index

Product Number: 5765-DME
5765-DAE

SC28-3173-02